

DOCUMENT RESUME

ED 086 722

TM 003 372

AUTHOR Yin, Robert K.
TITLE The Workshop and the World: Toward an Assessment of the Children's Television Workshop.
INSTITUTION Rand Corp., Washington, D.C.
SPONS AGENCY John and Mary R. Markle Foundation, New York, N.Y.
REPORT NO R-1400-MF
PUB DATE Oct 73
NOTE 61p.

EDRS PRICE MF-\$0.65 HC-\$3.29
DESCRIPTORS *Educational Assessment; Educational Innovation; *Educational Television; *Evaluation Methods; *Evaluation Needs; Instructional Television; *Program Evaluation

IDENTIFIERS *Childrens Television Workshop; Sesame Street; The Electric Company

ABSTRACT

The potential social impact of the Children's Television Workshop (CTW), originator of Sesame Street, the Electric Company, and other innovative educational ventures, and the fact that it is supported largely by public monies are strong reasons for examining the full implications of the workshop's achievements. This study was designed to identify the issues and potential measures for assessing the overall impact of the CTW, by determining whether CTW's impact can be captured within a comprehensive framework despite the diversity of the workshop's activities; the usefulness of existing data for such an assessment was examined. The study found that although each impact may be measurable, no acceptable method currently exists for uniting these separate impacts and thereby reaching a single statement of CTW's total impact. Thus, the study concluded that a comprehensive assessment of CTW would probably be a fruitless endeavor. The study recommended several independent analyses, including a national survey, a multi-year field study, a special investigation of the effects of Sesame Street and other CTW productions, an institutional study, a school-based field study, and a comparative study. Although the information obtained from such studies would not be agreeable into a single statement of CTW's effect, such a new array of evidence would substantially broaden the present understanding of CTW's accomplishments. (NE)

FILMED FROM BEST AVAILABLE COPY

THE WORKSHOP AND THE WORLD: TOWARD AN ASSESSMENT OF THE CHILDREN'S TELEVISION WORKSHOP

PREPARED UNDER A GRANT FROM THE JOHN AND MARY R. MARKLE FOUNDATION

ROBERT K. YIN

R-1400-MF

OCTOBER 1973

ED 086772

25th
Year

Rand
SANTA MONICA, CA 90406

The research described in this report was supported by a grant from The John and Mary R. Markle Foundation. Reports of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

The research reported in this publication was done in the Washington Office of The Rand Corporation, 2100 M Street, N.W., Washington, D.C. 20037

Published by The Rand Corporation

Cover photograph courtesy of the Children's Television Workshop.

ED 086772

THE WORKSHOP AND THE WORLD: TOWARD AN ASSESSMENT OF THE CHILDREN'S TELEVISION WORKSHOP

PREPARED UNDER A GRANT FROM THE JOHN AND MARY R. MARKLE FOUNDATION

ROBERT K. YIN

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

TM 003 322

R-1400-MF
OCTOBER 1973

Rand
SANTA MONICA, CA. 90406

PREFACE

This study was undertaken because of the increasing diversity of the Children's Television Workshop (CTW). Such diversity has raised the question of how the impact and effectiveness of CTW's numerous activities might be assessed. Since CTW itself has been very sensitive to this question, the study was carried out with the full and enthusiastic cooperation of CTW staff members.

The reader should be aware of several important factors that provided the context for this study. First, CTW has encouraged numerous observers to study and write about its activities, and this has created a large body of documents and reports. Within the time limitations of this study, it was not possible to examine these documents exhaustively. Second, personal consultations with individuals knowledgeable about CTW formed an important source of information. Most of these consultations took place during July 1973, and many persons were on vacation or otherwise unavailable. Thus the final list of people consulted is in no way meant to be complete or even balanced. (A list of the individuals is found in the Appendix.)

Third, CTW has become a subject about which most people in public broadcasting and educational circles now have strong opinions, if not official positions. CTW's role in early childhood education, its role in the public broadcasting sector, and its gradual shift toward revenue-raising enterprises have become issues for discussion. For some persons, CTW represents a dream that has more than lived up to expectations; for others, CTW represents an educational venture that, while innovative, must not unjustly overshadow other innovations.

Given these factors, the study is nevertheless intended to present the major issues, ground rules, and recommendations for a potential future evaluation of CTW. Both the sponsors of CTW and its staff will, hopefully, find useful information concerning the CTW organization and the problems involved in assessing its activities.

SUMMARY

Sesame Street and *The Electric Company* undertake to entertain and simultaneously to teach specific cognitive skills to large television audiences. The two shows are produced by an innovative organization, the Children's Television Workshop (CTW), that seeks to use the mass media for educational purposes. Through these two popular television shows, CTW has already combined educational ventures with high-quality television entertainment; through newer ventures, including other television shows, CTW hopes to add to this accomplishment.

CTW's potential social impact at home and abroad, and the fact that CTW is supported largely by public monies, both provide strong reasons for examining the full implications of CTW's achievements. The purpose of this study is to identify the issues and potential measures for assessing the overall impact of CTW. The main goal of the study is to determine whether CTW's impact can be captured within a single, comprehensive framework in spite of the considerable diversity of CTW's activities. In addition, the study examines the usefulness of existing data on CTW for such an assessment.

Section I of the study begins by describing the CTW organization and its activities: the production of major television shows (*Sesame Street*, *The Electric Company*, and new shows); research and community education services; and other related activities (development of puppets and games, educational materials, foreign sales and foreign language production, cable television and television cassettes, and a feature-length movie). Section II then suggests that one framework for assessing the impact of these activities is to measure them in terms of organizational goals. Three illustrative goals are applied to CTW: to educate through television and related media, to create institutional innovation, and to raise revenue from television-related ventures. Each CTW activity can have one or more of these goals. The goals provide guidelines for measuring the impact of CTW's various activities: educational gains made by specified target and nontarget audiences (educational goal); broadcast time taken by new commercial and public television programs tailored after CTW's programs, number of classrooms using television as a curriculum supplement, number of research teams like that of CTW, and number of children served by CTW's viewing centers (all deal with the institutional innovation goal); and revenue (revenue-raising goal).

Two problems arise in considering these measures. First, the measures do not reflect the social value of the impact. It is difficult, for instance, to determine how much a statistically measured cognitive gain is worth in societal terms. Second,

critics have cited several potentially ambivalent impacts of CTW: the untoward effects of excessive television viewing; the use by commercial networks of institutional innovations such as CTW's "distractor" technique for promoting programs with potentially adverse effects; and the possible aggravation, in producing cognitive gains among the general population, of an achievement disparity between advantaged and disadvantaged children. The social value of these ambivalent impacts, if any, is also difficult to determine, and even if determinable cannot be readily deducted from the positive impacts of CTW. These two problems pose a serious obstacle to any attempt to reduce CTW's activities to a single, summary impact statement.

Because CTW has been the subject of many prior studies, a large body of data on its activities already exists. Section III of the study reviews these data for their potential usefulness for future assessments of CTW. Included in the review are studies by Herbert Sprigle, Herman Land, public opinion and audience measuring organizations, the Educational Testing Service, and Thomas Cook. The review reaches the following conclusions: None of the existing studies can serve as a summary evaluation of CTW, since they focus mainly on the impact on the target audiences of *Sesame Street* and *The Electric Company*, and do not cover nontarget audiences or CTW's other activities; the public and audience opinion surveys need to be repeated under more rigorous conditions in order to determine the characteristics of the national viewing audience; and the ETS evaluations that provide adequate data on cognitive gains are limited to a few sites, and therefore cannot be generalized to a national population. The existing data, in short, do not easily lend themselves to use as part of an overall assessment of CTW. Section III also concludes, however, that many of the problems will still remain even in a totally new evaluation. Most important, a new and comprehensive evaluation of CTW is likely to raise more issues than it will settle. Another difficulty is that, because of the apparent high viewer rates, any field studies of *Sesame Street* or *The Electric Company* will probably not have control groups amenable to generalizations about national populations.

Section IV then raises another important issue for any further assessment of CTW: the purpose of the assessment and its potential audience. The potential audiences are national sponsors of children's television programs (e.g., the U.S. Office of Education, the Corporation for Public Broadcasting, and the foundations), or CTW itself. An assessment aimed at the former would attempt to guide resource allocation decisions among different television programs; an assessment aimed at the latter would guide internal managerial decisions. Both the purpose and audience must be determined before any assessment is actually conducted. This is because the purpose and audience will dictate the justifiable costs as well as the issues to be studied.

Section IV further discusses the major problem in creating any summary of CTW's activities, no matter what the purpose or who the audience. This is a problem of aggregating diverse impacts, some positive and others negative. Even a single program such as *Sesame Street* has at least the following impacts:

- Cognitive effects among the target population;
- Cognitive effects among nontarget populations;
- Innovation effects (e.g., number of new network programs that follow the same format or goals; incremental use of television in classrooms due to *Sesame Street* innovation);

- Revenue raised by foreign sales or production of the program;
- Possible effects of decreased attention span and possible effects of subsequent overinfluence by television; and
- Possible effects due to aggravation of the educational gap between children of low- and middle-income families.

Although each impact may be measurable, no acceptable method currently exists for aggregating these separate impacts, much less for combining the overall impact of *Sesame Street* with that of other CTW activities.

For this major reason, *the study concludes that a comprehensive assessment of CTW, attempting to combine all of its activities into the same analytical framework, will probably not be a fruitful endeavor. Rather, the study recommends several independent analyses, not necessarily to be carried out by the same research teams or with the same sources of financial support.* These independent analyses would include the following six studies that could fill important gaps in the existing knowledge about CTW's activities:

- A national survey, conducted by a research rather than a polling organization, that would determine the full size and characteristics of the audiences (target and nontarget) of CTW's programs, especially in comparison to those of other television programs;
- A multi-year field study, with specially designed experimental and control groups, of the cognitive effects of viewing *Sesame Street* and *The Electric Company* alone and in sequential combination;
- A special investigation of the effects of *Sesame Street* and *The Electric Company* on nontarget population audiences; in particular, the effect of *Sesame Street* on slow-learning children and foreign-speaking adults, and the effect of *The Electric Company* on preschool children;
- An institutional study examining the revenue-raising potential of activities in two or more nonprofit education organizations, including CTW;
- A field study to determine the actual amount of in-school effort required to teach the same skills as taught by *Sesame Street* and *The Electric Company*; and
- A comparative study of the costs and effects of three preschool education programs, *Sesame Street*, *Misterogers' Neighborhood*, and *Captain Kangaroo*, covering both cognitive and noncognitive aspects.

These six studies represent the potential next steps for further assessing CTW. The general goal in undertaking these studies is to increase knowledge about the diversity of CTW's effects. Even though the information will not be available into a single statement of CTW's impact, such a new array of evidence, combined with the existing evaluations of CTW, will add substantially to our understanding of the impact of CTW on the world.

ACKNOWLEDGMENTS

This report would not have been possible without the collaboration of the people whom I contacted during the summer of 1973. The Appendix lists these people, and I am grateful for their assistance.

Several individuals were also helpful in commenting on earlier drafts of the report, and I would like to thank them in addition: Robert Davidson, Robert Hatch, Thomas Kennedy, and Edward Palmer of CTW, and David Armor, Stephen Barro, George Comstock, Henry Geller, John Pincus, and Senta Raizen of The Rand Corporation. Finally, Beth Hunt of Rand made the task much easier with accurate and speedy typing assistance on earlier drafts. None of these people, of course, bears any responsibility for the final product.

CONTENTS

PREFACE.....	iii
SUMMARY	v
ACKNOWLEDGMENTS.....	ix
Section	
I. THE CHILDREN'S TELEVISION WORKSHOP.....	1
Television and Children.....	2
The Workshop Organization.....	5
II. MEASURING THE IMPACT OF THE CHILDREN'S TELEVISION WORKSHOP	10
Impact Measured in Terms of Goals.....	11
What To Measure	14
Educational Impact.....	14
Impact on Institutional Innovation	19
Broadcast Innovation.....	19
Education Innovation	20
Research Innovation	20
Urban Neighborhood Innovation	20
General Guidelines	20
Financial Impact.....	22
Summary of Goal-Oriented Impact Measures	22
Measuring the Potential Negative Impact of CTW Activities	22
Summary	24
III. EXISTING DATA ON THE CHILDREN'S TELEVISION WORKSHOP.....	26
Sprigle Study	27
Land Study	27
Audience Surveys	28
First ETS Evaluation of Sesame Street	29
Second ETS Evaluation of Sesame Street.....	31
Third ETS Study: Evaluation of The Electric Company	32

Cook's Reanalysis of the ETS Data	34
Summary of Existing Evaluations	34
Potential Problems for a New Evaluation	36
 IV. CONSIDERING A NEW ASSESSMENT OF THE CHILDREN'S TELEVISION WORKSHOP	37
Possible Purposes for an Evaluation and Its Potential	
Audiences	37
Impact Evaluation	38
Strategy Evaluation	39
Project Evaluation	39
Implications of Alternative Evaluations	40
Problems in Aggregating the Impact of CTW's Activities	41
Summary	44
 V. CONCLUSIONS AND NEXT STEPS	45
 APPENDIX: INDIVIDUALS CONSULTED	49

I. THE CHILDREN'S TELEVISION WORKSHOP

Sesame Street became a household word with the advent of the television program of the same name in 1969. The program now claims to reach nine million of the nation's twelve million children who are two to five years old. More importantly, it undertakes to teach specific cognitive skills to these children. *Sesame Street* thus stands as a demonstration that a television program with the goal of instructing preschool children can also be entertaining. The demonstration is based on formal educational evaluations and on the very personal reactions of parents who have observed directly the educational effects of *Sesame Street* on their children. Even in 1970, when one-half of the population still could not receive public television programs, *Sesame Street* was already the most frequently mentioned program when parents were asked, "Which is the best program for your child?" (See Table 1.)

One direct result of *Sesame Street*'s popular success was the creation of a second television program, *The Electric Company*, which began broadcasting in 1971. It attempts to entertain children in the seven to ten year age group while at the same time teaching them basic reading skills. As with *Sesame Street*, the effects of *The Electric Company* have also been assessed by formal educational evaluations. While

Table 1

PROGRAMS MOST FREQUENTLY MENTIONED BY PARENTS AS BEING
THE BEST FOR THEIR CHILDREN

1960			1970		
Program	N	%	Program	N	%
Captain Kangaroo	217	19	Sesame Street	220	30
Lassie	178	15	Walt Disney	167	23
Walt Disney	144	12	Lassie	80	11
Romper Room	115	10	Captain Kangaroo	68	9
Father Knows Best	115	10	Wild Kingdom	53	7
Huckleberry Hound	78	7	Jacques Cousteau	53	7
Popeye	71	6	Romper Room	42	6
Dennis the Menace	65	6	National Geographic Specials	30	4

SOURCE: Robert T. Bower, *Television and the Public*, Holt, Rinehart and Winston, New York, 1973, p. 164.

The Electric Company, because it was created second, may not achieve the same fame as *Sesame Street*, the show nevertheless has also been quite popular. A survey of in-school viewing during the first two months after the show went on the air, for instance, found that classes in 23 percent of all elementary schools in the United States had already adopted the program; this figure rises to 37 percent if schools without television sets are not counted.¹

TELEVISION AND CHILDREN

Both of these shows were created during a period of increasing public concern over the effects of television on children, and both are primarily supported by federal funds. Television, of course, has become the dominant source of all public information. The continued rise in importance of the medium is shown by a recent longitudinal survey (see Table 2). Moreover, because Americans have traditionally been sensitive to the long-term effects (intended and unintended) of early childhood experiences, the subject of television's impact on children touches upon deep economic and emotional nerves. On the economic side, television programs designed specifically for children provide an important advertising outlet for some industries and an important source of income for the television networks.² On the emotional side, most people who have young children have strong opinions about children's television. Because these opinions are based on a person's role as a parent, the views tend to be well formed and emotionally charged.

Thus it is not surprising that children's television continues to draw public attention and debate. For example:

- The effect of television on children has twice been the subject of national inquiry, first as part of a larger study by the National Commission on the Causes and Prevention of Violence (1969), which reviewed past research on the relationship between violence viewing and aggression, and then by a specially commissioned Surgeon General's study on television and social behavior (1971), which reviewed past research and sponsored new research.³ Both inquiries have stimulated, but not settled, arguments over the effect of television violence on the subsequent behavior of children.
- The role of advertising in children's television programs has been a controversial matter, with parties such as Action for Children's Television (ACT) recommending the complete abolition of such advertising. The precise form of regulatory action by the federal government is still being considered by the Federal Communications Commission.⁴

¹ Robert E. Herriott and Roland J. Liebert, "The Electric Company In-School Utilization Study," Institute for Social Research, Florida State University, August 15, 1973.

² A recent analysis of the economic aspects of television programs designed specifically for children is presented in Alan Pearce, "The Economics of Network Children's Television Programming," mimeographed paper, Federal Communications Commission, Washington, D.C., July 1972.

³ National Commission on the Causes and Prevention of Violence, *Mass Media and Violence*, Vol. IX, U.S. Government Printing Office, Washington, D.C., 1969; and *Television and Social Behavior*, five volumes, U.S. Government Printing Office, Washington, D.C., 1972.

⁴ See Docket No. 19142, in the matter of "Petition of Action for Children's Television (ACT) for rule making looking toward the elimination of sponsorship and commercial content in children's programming and the establishment of a weekly fourteen hour quota of children's television programs," Notice

Table 2
OPINIONS ABOUT THE MASS MEDIA
(Percent)^a

Survey Question	Television		Magazines		Newspapers		Radio		None/NA	
	1960	1970	1960	1970	1960	1970	1960	1970	1960	1970
Which of the media:										
Is the most entertaining?	68	72	9	5	13	9	9	14	1	0
Gives the most complete news coverage?	19	41	3	4	59	39	18	14	1	2
Presents things most intelligently?	27	38	27	18	33	28	8	9	5	8
Is the most educational?	32	46	31	20	31	26	3	4	3	5
Brings you the latest news most quickly?	36	54	0	0	5	6	57	39	2	1
Does the most for the public?	34	48	3	2	44	28	11	13	8	10
Seems to be getting worse all the time?	24	41	17	18	10	14	14	5	35	22
Presents the fairest, most unbiased news?	29	33	9	9	31	23	22	19	9	16
Is the least important to you?	15	13	49	53	7	9	15	20	7	5
Creates the most interest in new things going on?	56	61	18	16	18	14	4	5	4	5
Does the least for the public?	13	10	47	50	5	7	12	13	23	20
Seems to be getting better all the time?	49	38	11	8	11	11	10	15	19	28
Gives you the clearest understanding of the candidates and issues in national elections?	42	59	10	8	36	21	5	3	7	9

SOURCE: Robert T. Bower, *Television and the Public*, Holt, Rinehart and Winston, New York, 1973, p. 14.

^a1960 base: 100 percent = 2,427;
1970 base: 100 percent = 1,900.

- Direct federal support for public television, including children's programs, has increased via the Corporation for Public Broadcasting (CPB), which contracts with producers for the distribution rights to programs. The continuing nature and extent of the government's role, however, is still open to question, and discussion is aggravated by the rivalry between the CPB and the Public Broadcasting Service (PBS).⁵
- Direct federal support for children's television programs has also increased under the sponsorship of the Department of Health, Education, and Welfare,

of Inquiry and Notice of Proposed Rule Making, released on January 26, 1971. The FCC heard oral arguments on this issue during the following year, and the debate has attracted far more individual inquiries and letters than any other FCC issue.

⁵ The document that called for the creation of the Corporation for Public Broadcasting is the Report of the Carnegie Commission on Educational Television, *Public Television: A Program for Action*, Bantam Books, New York, 1967. CPB was founded in 1968; in March 1970 it created PBS, and since that time the division of responsibilities between the two organizations, especially in controlling new programs, has been unclear.

and in particular, the Office of Education. This support goes not only to public broadcasting programs but also to the production of special segments on commercial programs. However, the rationale for this supportive role has not been clearly established, and thus comes under occasional scrutiny. For instance, during the summer of 1973, staff members of the House Committee on Appropriations examined individual project grants, and indirectly touched upon policies underlying governmental support for children's programs.

That *Sesame Street* and *The Electric Company* were created during this period of heightened public awareness of the effects of television on children is no accident. The original plans for *Sesame Street*, as conceptualized by Joan Ganz Cooney, began with the observations that television was already an important part of the child's early life, but that little had been done to utilize television as a positive educational experience.⁶ *Sesame Street* thus represented a deliberate attempt to deal positively and creatively with the issue of television's educational impact on children. Whereas the more common approaches to this issue had been and continue to be negative in that they attempt to suppress the child's exposure to undesirable television programs, *Sesame Street* and then *The Electric Company* took on the far more creative task of showing how new programs could be developed and successfully aired.⁷

Because of their unique accomplishments, both *Sesame Street* and *The Electric Company* deserve careful analysis to determine the extent and nature of their social impact. At stake in this analysis are several major policy issues:

- First, the experiences of *Sesame Street* and *The Electric Company* should shed light on the effectiveness of using the television medium for mass education.
- Second, the experiences should help to determine the differences between using public as opposed to commercial broadcasting distribution for such programs.
- Third, the outcome of any analysis will provide more information on the rationale (or possible lack of it) for the government's role in supporting children's television programs, whether broadcast over the public or commercial systems.

Carrying out such an analysis, however, is not an easy task. Many important factors cannot be adequately measured. Even among those that can be measured, there are a variety of direct and indirect effects to be assessed. These include: changes in the cognitive and noncognitive skills among the target population of children; changes in other attitudes and behaviors, particularly toward television or education, among those children; changes among the parents and other family members of the intended target populations; changes among teachers or schools; changes among children of different ages from the target population; and changes

⁶ These views are reflected in two documents that laid out the initial rationale and plans for developing new programs: Joan Ganz Cooney, "The Potential Uses of Television in Preschool Education: A Report to the Carnegie Corporation of New York," October 1966; and Joan Ganz Cooney and Linda Gottlieb, "Television for Preschool Children: A Proposal," Carnegie Corporation of New York, February 19, 1968.

⁷ That the programs were filling a void is quite clear. For instance, the authors of one of the early studies of television's impact had noted that: "Concerning the cognitive effects of television, the general conclusion is one of disappointment. This is not because television is doing any special harm in this respect, but rather because it isn't realizing its full potential as a carrier of ideas and information." See Wilbur Schramm, et al., *Television in the Lives of Our Children*, Stanford University Press, Stanford, California, 1961, p. 173.

in the broadcasting industry and production of other programs, particularly those intended for children.

The purpose of the following study is to begin the groundwork for an evaluation. The study will focus on the possible measures to be used, and will review existing data on *Sesame Street* and *The Electric Company* for their adequacy. The study will also raise questions concerning the possible uses and audience for such an evaluation, and recommend the next steps to be taken.

THE WORKSHOP ORGANIZATION

The broader context for analyzing the accomplishments of *Sesame Street* and *The Electric Company* is the Children's Television Workshop (CTW), an organization that includes a number of activities related to the production of these two programs. The effect of the television programs can best be understood in light of this organization for two reasons. First, any generalizations about the accomplishments of the two television programs, especially if used in order to develop new programs, must take into account the organizational ability to plan and produce new programs. Second, the Children's Television Workshop also carries out other activities, such as community education services, that are clear attempts to reinforce the impact of the two television programs, and that therefore also need to be examined.

CTW was founded in 1968. Its initial activity was the production of *Sesame Street*, and these efforts required not only a television production staff, but also two elements that have since become part of the distinctive CTW style:

- **Research**, to help identify educational goals, develop television segments incorporating these goals, and evaluate the effectiveness of the programs in achieving the goals; and
- **Community education services**, a trained staff of field personnel who encourage and reinforce viewing of the television programs through the establishment of neighborhood viewing centers and other activities in low-income areas. The centers often provide the initial opportunity for children to view the programs, and distribute parents' guides and other materials related to the programs.

Since the first airing of *Sesame Street* in November 1969, CTW has gradually expanded and diversified its activities. These include the second television program, *The Electric Company*, first broadcast in October 1971, and several related television activities:

- **Foreign sales**, which promote the broadcasting of *Sesame Street* in other countries. The first programs broadcast outside the United States were broadcast in Central America in 1970; since then, *Sesame Street* has been shown in Africa, East Asia, Europe, and the Middle East.
- **Foreign language production**, which has developed three new versions of *Sesame Street* that were produced entirely in foreign countries: *Plaza Sesamo* (Spanish), *Sesamstrasse* (German), and *Vila Sesamo* (Portuguese).
- **New television programs**, with the major effort currently going into the

development of a health show, scheduled to begin broadcasting in 1974, and with substantial planning efforts for a history show and a children's show emphasizing affective learning.

CTW has also developed a number of nonbroadcasting activities. These activities concern further advancements in either the education or television fields, but do not necessarily involve the production of new television programs, and include:

- **Nonbroadcast products**, which is the oldest nonbroadcast activity and which has been responsible for licensing manufacturers to produce program-related puppets, games, puzzles, books, records, and children's magazines. There are now about 100 such products, and many are marketed in both the United States and overseas.
- **Educational materials**, a relatively new venture designed to produce multimedia materials for elementary schools. The first such materials will be sold during 1974.
- **New investments**, or the development of innovative means for the use of media in educating children. These presently include interests in cable television franchising and programming, the use of television cassettes for public education films, and the development of a full-length feature film.

The evolution of CTW's budget through FY 1974 partially reflects its involvement in these activities (see Table 3). For purposes of further analysis, however, all these activities will be broken into twelve separate items:

1. Sesame Street
2. The Electric Company
3. Health show and other new shows
4. Community education services
5. Research
6. Nonbroadcast products
7. Educational materials
8. Foreign sales
9. Foreign language production
10. Cable TV
11. TV cassettes
12. Movie

Several previous reports have already traced the early development of CTW.⁸ These reports make clear that CTW's accomplishments have demonstrated the following:

⁸ These reports include: Richard M. Polsky, "The Children's Television Workshop, 1966-1968," D.Ed. dissertation, Teacher's College, 1973; Herman W. Land, *The Children's Television Workshop: How and Why It Works*, Nassau Board of Cooperative Educational Services, Jericho, New York, 1972; Gerald S. Lesser, "Learning, Teaching, and Television Production for Children," *Harvard Educational Review*, Vol. 42, May 1972, pp. 232-272; and Gerald S. Lesser, "Children and Television: Lessons from Sesame Street," Harvard University Graduate School of Education, unpublished manuscript, January 1972. There are numerous popular descriptions of *Sesame Street* and the CTW organization. Two of the more comprehensive are: Phyllis Feinstein, *All about Sesame Street*, Tower Publications, New York, 1971; and Martin Mayer, *About Television*, Harper and Row, New York, 1972, Chapter six.

Table 3

CTW EXPENDITURES, BY ACTIVITY, FY 1969-1974

Activity	Fiscal Year (\$000)					
	Actual 1969	Actual 1970	Actual 1971	Actual 1972	Actual 1973	Budget 1974
Sesame Street	\$ 918.0	\$4,723.0	\$4,538.1	\$ 4,309.1	\$ 4,660.3	\$ 4,771.9
The Electric Company	--	--	1,000.5	5,332.7	5,193.0	4,825.4
Health show	--	--	--	--	313.4	725.0
Commercial stations	--	--	47.2	95.5	142.4	100.0
Community education services	119.7	641.3	671.2	1,344.2	1,255.7	1,076.9
Broadcast administration	156.8	300.6	432.8	594.3	747.7	780.2
Total, Broadcast	1,194.5	5,664.9	6,689.8	11,675.8	12,312.5	12,279.4
Nonbroadcast products	--	--	1,077.9	1,747.4	2,214.7	3,526.1
Foreign sales division	--	--	496.6	539.9	818.8	786.4
Foreign language production	--	--	119.5	533.0	1,685.6	890.4
Total, Non- broadcast	--	--	1,694.0	3,042.7	4,719.1	5,202.9
Research	(176.3)	(429.3)	(601.8)	(865.8)	(869.4)	(872.2)
Public information	(99.7)	(290.7)	(510.4)	(851.5)	(705.2)	(743.4)
Corporation services	(271.5)	(472.8)	(787.5)	(1,503.0)	(1,516.9)	(1,875.6)
Total Corporate Support	(547.5)	(1,192.8)	(1,899.7)	(3,220.3)	(3,091.5)	(3,491.2)
Special purpose grants	--	--	103.0	67.3	118.4	26.4
Other expenses	--	--	134.5	185.0	--	--
Total Expenses	1,194.5	5,664.9	8,621.3	14,748.4	17,150.0	17,508.7
Investment fund activities ^a	--	--	--	--	454.2	653.8
Capital expenditures ^b	--	--	178.0	220.1	222.1	210.2
GRAND TOTAL	\$1,194.5	\$5,664.9	\$8,799.3	\$14,968.5	\$17,826.3	\$18,372.7

^aIncludes activities such as cable television franchising, production of full-length film, and development of materials for television cassettes.

^bCapital expenditures are not included in CTW expenses as they are carried, net of depreciation on CTW's balance sheet.

- Television can be used as a medium that teaches cognitive skills while it simultaneously entertains an extremely large audience,
- Children's television can be effectively designed for viewing by specific age groups *within* the two to twelve year range (the entire range was previously used by commercial television as the only definition of the child-viewing audience);
- Television can penetrate a large majority of homes using public, and not commercial broadcasting stations; and
- Television production and research can be combined to produce more effective programs and to evaluate the impact of those programs.

Because its basic goal is the creative use of the mass media for educational purposes, CTW is a nonprofit, or tax-exempt institution. The primary financial support for CTW comes from the federal government and foundations, although revenue from nonbroadcasting activities is also used to support the television programs (see Table 4).⁹ The sources and amount of support have been issues of serious management concern, with the natural uncertainties of annual funding cycles compounded by increased pressure from the federal government and the foundations: (1) to support newer ventures, if any, rather than continuing programs like *Sesame Street*, and (2) to rely more heavily on revenue-raising activities as a source of income.

This brief overview of CTW's activities sets the scene for assessing CTW's impact. Section II attempts to conceptualize a series of relevant measures; Section III reviews existing evaluations of CTW and examines their continuing usefulness; Section IV raises the major issues to be settled before attempting a new evaluation; and Section V recommends the next steps to be taken.

⁹ For FY 1974, the estimated expenditures exceed income. These expenditures will be reduced if additional income is not forthcoming.

Table 4
CTW INCOME, BY SOURCE, FY 1969-1974

Source	Fiscal Year (\$000)					
	Actual 1969	Actual 1970	Actual 1971	Actual 1972	Actual 1973	Budget 1974
Office of Education Corporation for Public Broadcasting	\$ 833.3 --	\$3,166.7 750.0	\$2,900.0 500.0	\$ 7,000.0 2,000.0	\$ 6,000.0 5,000.0	\$ 3,000.0 5,200.0
Total, Government	833.3	3,916.7	3,400.0	9,000.0	11,000.0	8,200.0
Ford Foundation	250.0	1,288.0	1,000.0	1,000.0	--	--
Carnegie Corporation	100.0	1,400.0	600.0	1,000.0	500.0	--
Markle Foundation	--	250.0	--	--	37.5	25.0
Mobil Corporation	--	--	--	250.0	--	--
Johnson Foundation	--	--	--	--	236.8	200.0
Exxon Foundation	--	--	--	--	37.5	200.0
Commonwealth Foundation	--	--	--	--	--	100.0
Commercial stations	--	--	115.0	241.0	303.4	225.0
Other	11.2	186.9	16.0	52.6	8.0	--
Total, Institutional	361.2	3,124.9	1,731.0	2,543.6	1,123.2	750.0
Nonbroadcast products	--	103.6	1,531.0	2,201.4	1,994.5	4,187.8
Foreign sales division	--	--	363.0	851.4	1,964.6	1,437.7
Foreign language production	--	--	--	260.0	1,711.0	975.0
Special purpose grants	--	--	103.0	67.3	112.0	19.6
Other revenue	--	31.3	55.0	60.9	71.5	90.0
Total, CTW	--	134.9	2,052.0	3,441.0	5,853.6	6,710.1
Total, Income	1,194.5	7,176.5	7,183.0	14,984.6	17,976.8	15,660.1
Investment fund grant from Ford Foundation ^a	--	--	--	--	2,000.0	--
GRAND TOTAL	\$1,194.5	\$7,176.5	\$7,183.0	\$14,984.6	\$19,976.8	\$15,660.1

^aFord made a multi-year grant in 1972 to CTW to provide venture capital funds. The grant totals \$6 million, and can be drawn in varying amounts over a seven-year period.

II. MEASURING THE IMPACT OF THE CHILDREN'S TELEVISION WORKSHOP

The activities of the Children's Television Workshop (CTW) have produced an impact on many different audiences. CTW has not only affected children's learning, but also may have changed parent and teacher attitudes toward early learning, influenced the use of television in schools, and changed attitudes toward producing new public broadcasting programs. Many of these effects are widely accepted, even though few attempts have been made to assess them precisely. Moreover, some of the effects may not be readily measurable.

Any assessment of CTW's impact runs directly into the problem that what is socially significant, i.e., accepted broadly as a major accomplishment, is often quite different from what is measurable. And what is currently measurable in no way captures the overall impact of a single television program, much less the work of an entire organization. On the matter of television's direct impact on children alone, for instance, Gerald Lesser has identified several potential benefits of television.¹⁰ Television may serve to provide children with enjoyment, a temporary sanctuary or refuge from daily life, an alternative means of learning, knowledge about the world as it is, and knowledge about the world as it might be. Social scientists currently have few ways of assessing these benefits, even with substantial evaluation efforts. Nor can the benefits from television's ability to serve special audiences not previously served be adequately measured.

Similarly, in assessing the impact of CTW on public broadcasting, one observer has remarked that public sponsors are now beginning to ask two questions of any new program: What are the program's goals? Do these goals serve important public needs? Such questions stem directly from the experiences of *Sesame Street* and *The Electric Company*, and would have been foreign to the world of public broadcasting five years ago. At that time, the major criterion for a successful public broadcasting program might have been the review of the program by television critics. Success or failure, in other words, was still defined by the television industry itself. What is important now, however, as a result of the new attention to audience needs, is that the *public* is finally becoming an integral part of the public broadcasting system. Few methods exist that can measure this change, or satisfactorily measure the equally important claim that *Sesame Street* has changed the public's expectation and image of the television medium.

¹⁰ Gerald S. Lesser, "Children and Television: Lessons from *Sesame Street*," Harvard University Graduate School of Education, unpublished manuscript, January 1972.

IMPACT MEASURED IN TERMS OF GOALS

In spite of these limitations on what can be measured, one possible starting point is to *gauge CTW's impact in terms of a set of organizational goals*. This would be consistent with the general notion that new educational technologies should primarily be assessed and compared according to their *effectiveness*, i.e., the degree to which they achieve their stated goals, rather than according to their *social benefits*, i.e., their impact in furthering societal goals.¹¹ While such a judgment unfortunately avoids the ultimate question of the social worthiness or benefits of an organization such as CTW, there are simply too few guidelines to address this issue. Evaluation must begin with program effectiveness.

Even the determination of CTW's goals, however, is not a simple matter. Discussions with people knowledgeable about CTW suggest a wide variety of goals, explicit and implicit, and no necessary consensus. For the purposes of assessing singly each of CTW's activities, three general goals may be taken as illustrative examples:¹²

- To educate through the use of television and related media;
- To create institutional innovation in the educational and broadcasting communities; and
- To raise revenue from commercial ventures to support the broadcasting activities of the organization.

While the last of these three goals is an instrumental one, it is nevertheless singled out because it may be the only goal of individual activities within CTW.

Each of CTW's twelve activities may be characterized, in an oversimplified manner to be sure, in terms of these three general goals. There are at least two illustrative configurations of goal structure, which shall be referred to as CTW-1 and CTW-2.

CTW-1 would be based on the assertion that all CTW activities should *share* the basic philosophy of innovatively using the mass media to educate and entertain, but that some activities would have the additional goal of raising revenue for CTW. The rationale underlying each of the activities might be as follows:

- The broadcast activities, including any new programs to be developed, would follow the pattern set by *Sesame Street* and *The Electric Company*: Specific educational goals would be set and assessed, and an attempt would also be made to assess the programs' innovative impact on the educational and broadcasting communities. This pattern, for instance, would be applied to the health show, a new history show, the cable television venture, foreign sales, and foreign language production.
- The nonbroadcast products, including all games, puppets, puzzles, and books, would also have specific educational goals and be evaluated accordingly. The

¹¹ See Lawrence P. Grayson, "Costs, Benefits, Effectiveness: Challenge to Educational Technology," *Science*, Vol. 175, March 1972, pp. 1216-1222. For a more general statement, see Harry P. Hatry, "Measuring the Effectiveness of Nondefense Public Programs," *Operations Research*, Vol. 18, September-October 1970, pp. 772-784.

¹² These goals evolved from the author's discussions with many different people (see Appendix). The goals are meant, however, mainly for illustrative purposes, and are not intended either as a definitive assessment or as the goals set by CTW itself. Another goal, for instance, might simply be *entertainment*.

major innovative impact would be sought within the toy and game industry, where, it would be hoped, manufacturers would become more sensitive to the needs for establishing educational goals and performing summative evaluations.

- The multi-media educational materials to be produced for use in the elementary schools would follow a pattern similar to that of the nonbroadcast products.
- The development of a new feature film, among the new ventures, would also have specific educational goals and evaluation, and the institutional innovation would be sought within the movie industry, where, for example, a new pattern might be set for films that entertain and educate, or for films that are designed for in-theatre viewing by specially organized groups of school children.¹³
- The development of television cassettes would follow a pattern similar to either the broadcast or the movie activities.
- The research and community education services activities could be viewed as having separate innovation goals (e.g., increasing the role of research in other children's programs, or developing a new neighborhood institution around the viewing center), or merely as having the same goals as the broadcast and non-broadcast activities.

In addition, several activities would also be judged by their ability to raise revenue: the cable television venture, foreign sales and foreign language production, the nonbroadcast products, the educational materials, the feature film, and the television cassette venture. Figure 1 illustrates the division of CTW activities according to the first possible configuration, CTW-1.

CTW-2 would be based on the assertion that not all of CTW's activities must share the basic education/entertainment goals, but that some activities might be justified merely in terms of their ability to raise revenue. Thus, in CTW-2, only the broadcast activities, limited primarily to the domestic television programs, would attempt to follow the *Sesame Street* and *The Electric Company* pattern, with specific educational goals to be set and evaluated, and with an attempt to assess their innovative impact. The goals for research and community education services would be unchanged, but all of the other activities, except for educational materials, would be judged primarily in terms of their revenue-raising capabilities. Although such activities would not have an educational or innovation goal, they would still have some quality control procedures to ensure that the educational impact was not negative or undesirable. The educational materials would retain all three goals since such materials are part of the formal educational system and cannot, as with the other nonbroadcast products, ignore educational goals. Figure 2 illustrates the division of CTW activities according to the second possible configuration, CTW-2.

Although these two abstract and oversimplified configurations of goal structure appear to differ only slightly, in fact they connote considerably different managerial directions for the CTW organization. These two configurations are given as examples of the relationship between goals and the measurement of CTW's impact. They are only meant to illustrate two of several possibilities. However, since CTW-2 may more closely reflect the current direction of CTW than CTW-1, for discussion purposes we assume that the dominant theme is the second configuration.

¹³ Note, for instance, that movie theaters are normally closed before noon, and special movie programs could be developed for morning viewing by school groups.

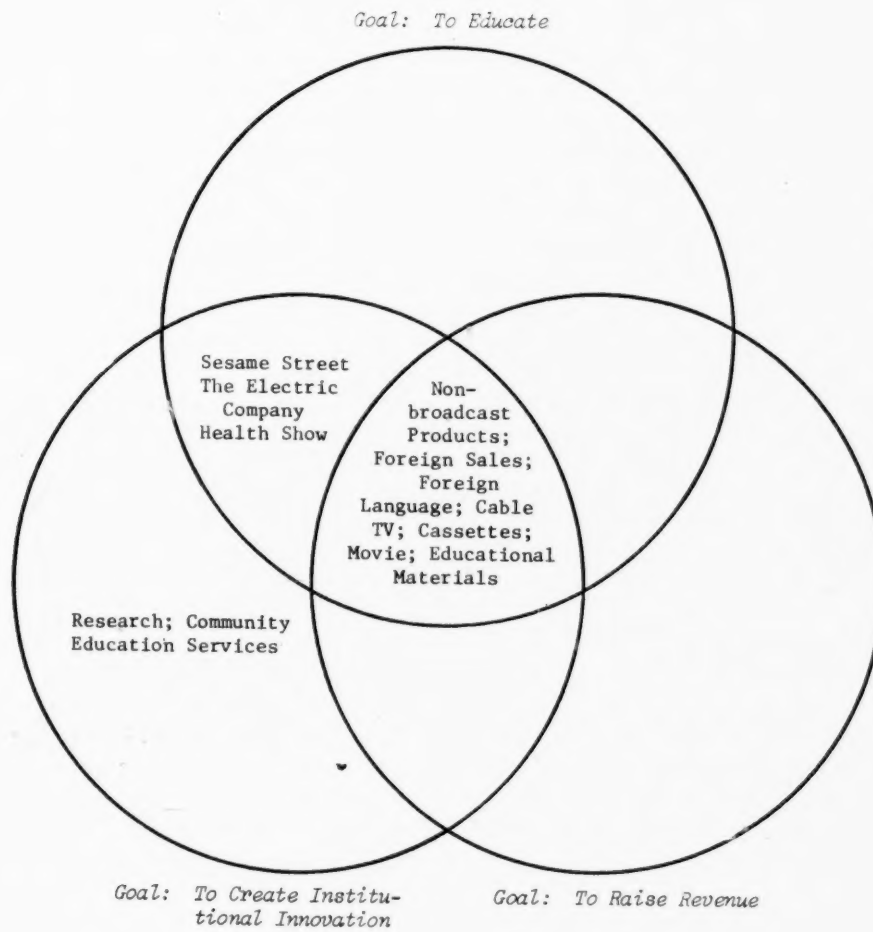


Fig. 1—One possible division of CTW activities according to three goals

Goal: To Educate

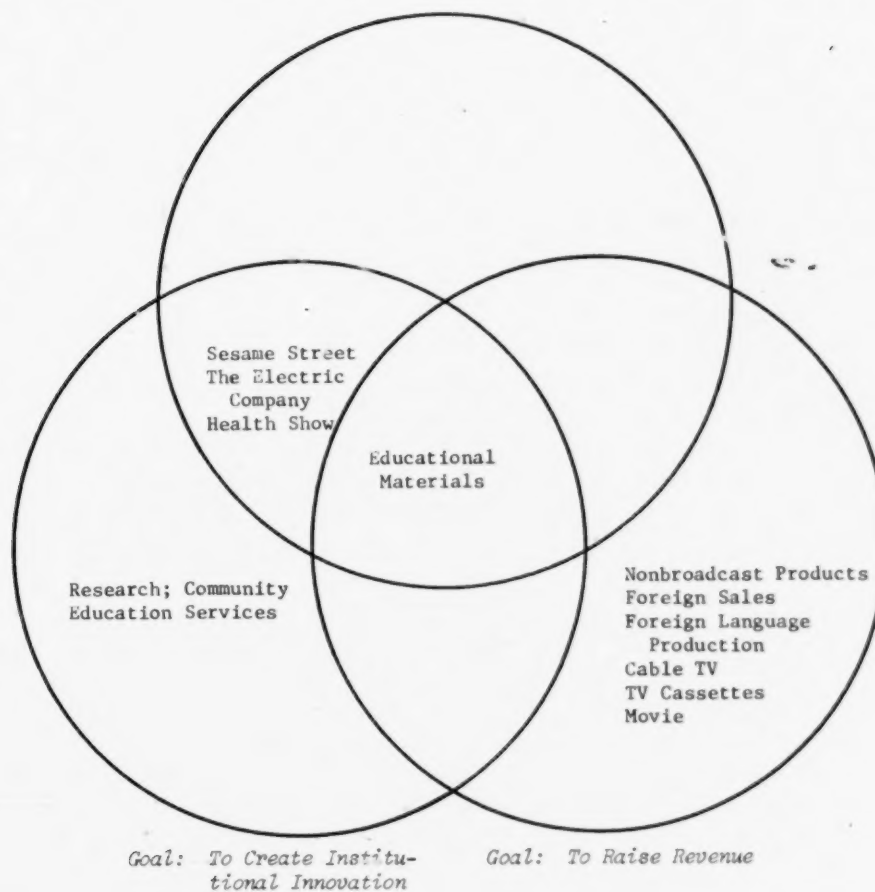


Fig. 2—Another possible division

WHAT TO MEASURE

The three goals suggest the main areas of measurement: educational impact, impact on institutional innovation, and financial impact. Regardless of how the CTW activities ultimately share in these three goals, specific measures might evolve in the following manner.

Educational Impact

The assessment of the educational impact depends in part on two questions: How are the effects of television programs to be measured? What constitutes a true measure of educational impact? In answer to the first, despite years of testing, there

are no standard methods for assessing the impact of television programs in general.¹⁴ Among children's programs, the dominant concern has been with the effects of televised violence. However, the behavioral consequences of viewing, whether by children or by adults, have not frequently been measured in traditional studies, which have focused on the subjects' verbal report (of prior behavior or of attitudes).¹⁵ In more recent studies where behavioral effects have been measured, the measurement has usually occurred under experimental conditions only, so that it is difficult to make policy-relevant statements concerning the overall impact of televised violence on aggression in American children. Thus, even though certain studies have developed useful measures of various types of children's behavior,¹⁶ there remains no analytic framework with which the impact of a specific television program can be assessed.

The second question relates to a deficiency in our conceptualization of educational programs, namely that the benefits of such programs have not been satisfactorily defined.¹⁷ For instance, suppose that an entirely new educational curriculum is developed and implemented for eighth grade junior high school students. The success of this curriculum can be judged by several different outcomes. The most important outcome from society's standpoint might be the demonstration of *long-term changes in the students' careers, income, and quality of life*. Yet much of the current research suggests that this "factory" view of the role of schools is misleading. Performance in the years following one's schooling is simply very poorly related to the common characteristics of schools. As an example, Christopher Jencks, et al., in their major study of educational inequality, found only weak relationships between educational curricula and subsequent career factors. The investigators concluded:¹⁸

It is true that schools have "inputs" and "outputs," and that one of their nominal purposes is to take human "raw material" (i.e., children) and convert it into something more "valuable" (i.e., employable adults). Our re-

¹⁴ There exist several major reviews of the impact of television on social behavior. See Wilbur Schramm, et al., *Television in the Lives of Our Children*, Stanford University Press, Stanford, California, 1961; Godwin C. Chu and Wilbur Schramm, "Learning from Television: What the Research Says," Stanford University, Institute for Communications Research, 1967; and Charles Atkin, et al., *Television and Social Behavior: An Annotated Bibliography of Research Focusing on Television's Impact on Children*, National Institute of Mental Health, Rockville, Maryland, PHS No. 2089, 1971.

¹⁵ For instance, see Gary A. Steiner, *The People Look at Television*, Alfred A. Knopf, New York, 1973, p. 82; and Wilbur Schramm, et al., *The People Look at Educational Television*, Stanford University Press, Stanford, California, 1963.

¹⁶ For instance, see Aletha Stein, et al., "Television Content and Young Children's Behavior," in *Television and Social Behavior, Vol. II: Television and Social Learning*, edited by John Murray, Eli Rubinstein, and George Comstock, U.S. Government Printing Office, Washington, D.C., 1972, pp. 202-317, for an interesting array of measures of "prosocial" behavior as applied to viewers and nonviewers of the television program *Misterogers' Neighborhood*.

¹⁷ This is true in spite of the existence of comprehensive handbooks on educational measurement. See, for instance, Benjamin S. Bloom, et al., *Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook I: Cognitive Domain*, Longmans, Green, New York, 1956. The most commonly used measure of an ultimate educational benefit is, not surprisingly, adult earning power. See Joint Economic Committee, *Benefit-Cost Analyses of Federal Programs*, 92d Congress, 2d Session, January 2, 1973, U.S. Government Printing Office, Washington, D.C., 1973.

¹⁸ Christopher Jencks, et al., *Inequality: A Reassessment of the Effect of Family and Schooling in America*, Basic Books, New York, 1972, p. 256. It should be noted, however, that the results of the Jencks study, as well as this major conclusion, have hardly been unequivocally accepted by other social scientists. See, for example, the "Review Symposium," including reviews by James S. Coleman, Thomas Pettigrew, William Sewell, and Thomas Pullum, in *American Journal of Sociology*, Vol. 78, May 1973, pp. 1523-1544.

search suggests, however, that the character of a school's output depends largely on a single input, namely the characteristics of the entering children. Everything else—the school budget, its policies, the characteristics of the teachers—is either secondary or completely irrelevant.

As a second alternative, the effect of any educational program might be examined in terms of *long-term changes in student performance within the school system*, and a second outcome of the hypothetical junior high curriculum might be the production of a larger number of high school or even college graduates. However, this second outcome has rarely been demonstrated, either. Few educational interventions in one school year, in other words, have shown lasting effects on school performance.¹⁹ One may interpret this shortcoming as an indicator of ineffective programs. Alternatively, one may posit a high depreciation factor in the carry-forward effects of grade-specific educational interventions. Thus it would be wrong and illogical to expect an intervention to produce a change beyond the immediately following time period. In fact, if a longitudinal effect on school performance is the criterion of success, all grade-specific interventions, taken singly, are likely to be judged failures. On the contrary, if there is a sound rationale for education intervention programs, it must rely on the cumulative results of major interventions during each year of schooling.²⁰

A third and narrower alternative for measuring the impact of a single education program, then, is merely to judge its impact on *student performance in the period immediately following the intervention program*. This measurement can be done in two ways: criterion-referenced testing that shows whether students have learned what the intervention program purports to teach, and achievement testing that shows whether the program has changed overall school performance. The hypothetical junior high program would be deemed successful, then, if the students participating in the program either retained some of what they learned or performed better at the outset of senior high school.

It is this third and last criterion that appears the most appropriate standard for assessing the impact of an educational intervention based on a single television program. Such a program, like *Sesame Street* or *The Electric Company*, can hardly be expected to produce a long-term impact, for instance, if in-school intervention programs have also failed in this respect. In fact, the demonstration of an early and immediate impact on student performance would actually be quite impressive. A recent and comprehensive review of educational research reached the major conclusion that:²¹

¹⁹ The longitudinal effects of the Head Start program and the whole question of evaluating early childhood education programs, of course, have been topics of heated controversy. For a start, see Westinghouse Learning Corporation and Ohio University, "The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development," July 12, 1969; Walter Williams and John W. Evans, "The Politics of Evaluation: The Case of Head Start," *The Annals*, Vol. 385, September 1969, pp. 118-132; Lois-ellin Datta, "A Report on Evaluation Studies of Head Start," U.S. Office of Child Development, 1969; Marshall Smith and Joan Bissell, "Report Analysis: The Impact of Head Start," *Harvard Educational Review*, Vol. 40, Winter 1970, pp. 51-104; and Victor G. Cicirelli, et al., "The Impact of Head Start: A Reply to the Report Analysis," *Harvard Educational Review*, Vol. 40, Winter 1970, pp. 105-129.

²⁰ This argument is similar to those made by Head Start evaluators. Smith and Bissell, for instance, state that "... in order to prevent cumulative and continuous retardation on the part of disadvantaged children, a policy of continued intervention during the elementary school years must supplement pre-school intervention programs." See Smith and Bissell, op. cit., p. 102.

²¹ Harvey A. Averch, et al., *How Effective Is Schooling? A Critical Review and Synthesis of Research Findings*, The Rand Corporation, R-956-PCSE/RC, March 1972.

Research has not identified any alternative educational practices, e.g., changes in school resources, process, organizations, and aggregate levels of funding, that are consistently related to students' educational outcomes.

For any single television program to show a significant short-run educational effect, in other words, would already be a stringent test of its educational impact.

While either the criterion-referenced tests or achievement tests may provide an adequate basis for assessing the educational effectiveness of CTW's activities, the results still need to be translated in order to determine the *social value* of the educational gains. The translation would allow the value of the gains to be compared with the value of achieving other goals. One of the few approaches to such a translation in any public service, including education, is to determine the prices that individuals would be willing to pay for the service.²² Because CTW purports to provide a supplemental educational experience that serves either a remedial or enriching purpose, one possible translation would be based on the present in-school costs (relying solely on traditional in-school methods) for producing the same educational gains as produced by the television programs. In other words, if evaluations show that *Sesame Street* produces a given amount of learning (by whatever test) among a certain number of children, a possible unit of measure of *Sesame Street's* educational impact would be the dollar cost of producing the same result via an in-school program that does not use television.²³

However, this translation is extremely difficult to make. Much further research would be needed to determine whether the in-school component, teaching the same skills as taught by *Sesame Street*, could even be isolated from other in-school activities. Furthermore, the translation would be more difficult in assessing the value of educational gains not traditionally associated with the formal educational curriculum. For instance, the value of CTW's new health show might, *in theory*, be measured in terms of the costs of an equivalent nontelevision effort to teach the same public health practices and information to the same number of people as will be reached by the health show.²⁴ There is little experience in the case of health information services, however, to show whether the public would actually support these costs.

In summary, the requisite information for assessing the educational impact of CTW's television programs would be:

²² See Julius Margolis, "The Demand for Urban Public Services," in Harvey S. Perloff and Lowdon Wingo, Jr. (eds.), *Issues in Urban Economics*, Johns Hopkins Press, Baltimore, Maryland, 1968, pp. 527-565.

²³ CTW occasionally has used this measure to make actual comparisons with in-school programs (for instance, see David D. Connell, "The Dollars and Sense of Sesame Street," *Dividend*, University of Michigan Graduate School of Business Administration, Winter 1971). What is suggested here, however, is that this unit be used to compare two or more television programs.

²⁴ There is some incipient confusion, incidentally, over the appropriate measures for CTW's new health show. Based on the author's recent study of telecommunications and health services (see William A. Lucas and Robert K. Yin, *Serving Local Needs with Telecommunications*, The Rand Corporation, R-1345-MF, November 1973), it would be unfair to hold specific telecommunications programs responsible for actual changes in health *status*. Changes in health status are a function of too many factors, both environmental and genetic, to be linked with specific intervention programs. Instead, the proper tests and measures should emphasize health *practices*. As an example, CTW's health show should be judged in terms of whether people, as a result of viewing, brush their teeth more frequently (assuming this to be one of the show's goals). It should not be judged in terms of whether the same people have fewer cavities.

- Data indicating the amount of educational gain, per viewer, on either a criterion-referenced test or an achievement test, or both;
- Data on the number of viewers of the television programs for a national population, and hence a rigorous estimate of the total number of viewers affected; and
- Estimates on how much the gains are worth, based on what the public would be willing to pay for similar gains in a nontelevision in-school program.

At the present time, the critical step in the assessment is the last one. New field research would have to establish the comparable efforts expended in in-school programs.

While the most important task would be to apply this approach to children who are within the target age groups for each television program, separate studies would have to deal with the impact on nontarget populations: e.g., children who are outside the targeted age groups but who may nevertheless watch the programs and show a measurable change in learning; slow-learning children; illiterate adults; and, in the overseas, gains not only in the learning of the mother tongue, but gains in learning English if the program is shown in English.

Note that none of these measures covers the potential noncognitive effects of the program. In fact, however, *Sesame Street's* curriculum has slowly incorporated a larger proportion of noncognitive goals, including affective and social aspects of behavior. The evaluation of achievements in the noncognitive area is a topic that has only begun to receive substantial research attention.²⁵ In preschool field studies, investigators have devised some behavioral measures and examined prosocial behavior in relation to television viewing. For instance, Stein, et al., observed the incidence of physical aggression, verbal aggression, cooperation, nurturance, verbalization of feelings, rule obedience, tolerance of delay, and task persistence among samples of preschool children.²⁶ Some of the children had viewed a diet of aggressive television programs as part of a summer preschool session; some had viewed *Misterogers' Neighborhood*; and some had viewed films of socially-neutral content. The results showed positive behavioral effects as a result of viewing a prosocial program like *Misterogers' Neighborhood*.

However, while isolated studies such as this one have produced important results, there are few standard measures of noncognitive behavior. Existing guidelines do not even suggest the range or limits to what should be measured and not measured. For instance, at a major attitudinal level, children's self-image and attitude toward learning may be improved by a program such as *Sesame Street*. How and whether this effect should be measured, and what the meaning of the results would be, are not yet part of the standard evaluation repertoire. The development of measures of noncognitive behavior may be seen as one of the challenges for the future.²⁷ At the present time, however, it may be worth focusing on the cognitive effects wherever possible, and considering noncognitive effects only in those cases, such as *Misterogers' Neighborhood*, where a program has noncognitive effects as its major goal.

²⁵ For a recent review, see Constance K. Kamii, "Evaluation of Learning in Preschool Education," in Benjamin S. Bloom, et al. (eds.), *Handbook on Formative and Summative Evaluation of Student Learning*, McGraw-Hill, New York, 1971, pp. 281-344.

²⁶ Aletha Stein, et al., op. cit., pp. 202-317.

²⁷ A current Rand study will in fact be addressing these questions in relation to a proposed evaluation emphasizing the noncognitive effects of Head Start programs.

Impact on Institutional Innovation

A second goal of CTW is to create change within certain broadly defined communities. For *Sesame Street* and *The Electric Company*, this includes the broadcast community, the educational community, the research community, and the urban neighborhood. If other CTW activities like nonbroadcast products or the full-length movie were included as having this goal, then other industries, e.g., toy manufacturers and the movie industry, would also be included.

The measurement of institutional innovation is only a crudely developed art. For one possible parallel, let us examine recent research on an institution that is somewhat like CTW, the *community development corporation* (CDC). There are several dozen such corporations scattered around the country, and at a broad organizational level, most have a similar goal as CTW: Most are concerned not with profit maximization, but with maximizing the community's welfare; and most are non-profit organizations with revenue-raising components.²⁸ The community welfare function, of course, is an entirely innovative one. CDCs are supposed to identify and solve just those very community problems that are important but that are not being dealt with by the existing private or public sectors.

The present state of research in evaluating CDCs, while very crude, suggests an approach for assessing CTW's impact on institutional innovation. This approach relies as much on *input* as on *output* measures.²⁹ Thus a CDC might be evaluated in input terms (e.g., the number of subscribers to a CDC-sponsored effort), and output terms (e.g., the number of community jobs it provides), or a mixture (e.g., the amount of capital flowing into the community).

However, the use of any of these measures still leaves open the question of the appropriate *value* given to these innovations. Without assessing the value, i.e., without translating each accomplishment to a common unit such as the prices people would pay for each innovation, the various accomplishments would again be difficult to interpret and compare. Although this valuation problem cannot be resolved, it is nevertheless worth reviewing the possible types of effectiveness measures.

Broadcast Innovation. There are several important and usable input measures for CTW's activities. *Sesame Street*, for instance, represents a new attempt to use television to educate children. It has several distinctive features, among them high-quality programming, integrated use of research, and clearly stated target populations and educational goals. Since the advent of *Sesame Street*, there have been several other television programs produced in the *Sesame Street* manner.³⁰ A crude measure of CTW's innovative impact on the television industry, then, would be to identify the amount of broadcast time taken by new television programs that

²⁸ The tension between revenue- and nonrevenue-raising activities within the CDC is in some ways similar to that of CTW. In the CDC, one social goal is providing jobs for high-risk unemployables; performance of this function clearly conflicts with profit maximization. In CTW, the desire to produce high-quality educational materials, to be sold for exceptionally low market prices in order to reach low-income consumers, may create a similar conflict.

²⁹ See, for instance, two recent CDC evaluation documents: Abt Associates, *An Evaluation of the Special Impact Program*, Phase I Report, March 1972, and Interim Report, March 1973; and Harvey A. Garn, et al., "CDC Evaluation," The Urban Institute, discussion paper, April 1973.

³⁰ According to one writer, CTW has served as a real impetus for better children's programs, and has had a greater impact than the FCC. See Timothy Green, *The Universal Eye: The World of Television*, Stein and Day, New York, 1972, p. 36.

clearly reflect the distinctive features of *Sesame Street* and *The Electric Company*.³¹ On commercial broadcasting, these could include *Multiplication Rock*, new segments inserted into the *Captain Kangaroo Show*, and other programs such as *Hot Dog* that may have since gone off the air.³² In public broadcasting, the measure would include the estimated increase in new children's programming supported by the Office of Education and other U.S. government agencies, which have increased support for such programs in part on the basis of the success of *Sesame Street*. New funds provided for educational television as part of the Emergency School Aid Act, for instance, were obtained by the Office of Education based on testimony about *Sesame Street*'s performance. A partial list of current grants for children's television programs is shown in Table 5.

Education Innovation. CTW's impact on education innovation primarily concerns the expanded use of television in the classroom. The impact may be assessed by identifying the number of classrooms that use television and for whom *Sesame Street* or *The Electric Company* had been the first or only viewing experience (this number would have to be adjusted by the projected rise, before CTW began broadcasting, of the use of TV in schools). A broader search might identify other curriculum changes due to the broadcasting of the two programs, and the programs' impact might then be measured by the amount of classroom time involved in these changes.

Research Innovation. CTW, through the use of the "distractor" technique³³ and its unique style of research, has also influenced research practices in educational television. Since very little program-specific research was carried out before CTW began, one measure of the impact of CTW's unique style of research could be the amount of new research used to produce and evaluate specific children's programs. Such research has been increasingly supported by the three major networks,³⁴ the Office of Education, the Office of Child Development, and the Corporation for Public Broadcasting, and might be measured in terms of the increased number of researchers now carrying out functions similar to those of CTW's research team.

Urban Neighborhood Innovation. Finally, CTW has also had a modest impact on urban neighborhoods, especially low-income communities. For instance, CTW has created a number of viewing centers in urban neighborhoods through its community education services. The impact of the centers could be crudely analyzed in terms of the number of children served or the number of hours each center is open. Further research might also identify ways of assessing the centers' impact in other community affairs.

General Guidelines. Whether analyzing only these activities or including some of the others such as nonbroadcast products, CTW's impact on institutional innovation may be measured by following the same steps for each activity:

³¹ Another distinctive feature of both shows is their portrayal of racial integration. The impact of this feature could also be traced in other children's programs, but this would be difficult to disengage from the effect of the country's generally increasing awareness of race and integration.

³² The use of this measure should not conceal the disappointment among the CTW staff that there has not been a more significant change on commercial television. Saturday morning children's programs, for instance, are much the same as they were before *Sesame Street*.

³³ The distractor technique, devised by CTW's Edward Palmer, assesses the interest in a television segment by measuring the frequency with which colored slides can distract a child viewing the segment.

³⁴ The National Broadcasting Company has even begun to use CTW's distractor technique.

Table 5

PARTIAL LIST OF CHILDREN'S TELEVISION PROJECTS
TO DEVELOP NEW PROGRAMMING^a

Project	Fiscal Year (\$000)						
	1968	1969	1970	1971	1972	1973	1974
1. Children's Television Workshop, New York City (OE) ^b	1,340	1,330	655	2,600	7,000	6,000	3,000
2. Appalachian Educational Lab., Charlestown, W. Va. (OE)					320	N.A.	
3. Rocky Mountain Educational Satellite Demonstration (OE)					255	1,224	
4. TV Serial on Infancy Nova University, Fort Lauderdale, Fla. (OE)						81	
5. Handicapped Segments for <i>Misterogers' Neighborhood</i> , Pittsburgh, Pa. (OE)					125	250	
6. Health Segments for <i>Captain Kangaroo</i> , Sutherland Learning Associates (OCD) ^c					250	250	
7. Career Segments for <i>Captain Kangaroo</i> , Sutherland Learning Associates (OE)					214	N.A.	
8. Bicultural Children's Television, Berkeley, California (OE-ESAA) ^d					890	3,500	
9. 3-4-5 Club, University of Houston, Tex. (OE)					300	N.A.	
10. Instruction through Television (Carrascolendes) Austin, Tex. (OE-ESAA)			215	260	537	1,269	
11. WNVN, Annandale, Va. (OE-ESAA)						1,762	
12. WTTW, Chicago, Illinois (OE-ESAA)						812	

^aExcludes projects dedicated to interconnection or other hardware development, teacher training projects, and parent training projects.

^bOE = Office of Education.

^cOCD = Office of Child Development.

^dESAA = A provision of the Emergency School Aid Act authorizes \$11.3 million for television programming grants in FY 1973 and \$6.8 million for FY 1974.

N.A. = Not available.

- Enumerate either the direct or indirect impacts of CTW activities, e.g., new programs or products tailored after a CTW program or product; and
- Measure the impact in terms related to the specific activity, e.g., the amount of broadcast time, the number of classrooms, or the number of children affected.

These measures should yield estimates of CTW's effectiveness in the various areas, and can ultimately be compared with the actual costs of CTW's efforts. However, any interpretation or comparison of the achievements will still require a translation, implicit or explicit, to some measure of the social value of the achievements.

Financial Impact

The impact in the third goal area, that of raising revenue to support CTW's broadcasting activities, is most easily measured. Here, the measure would simply be the amount of revenue that a given activity produces.

Discussion of the financial impact, however, does raise some matters that still seem unsettled within CTW. First, it is not clear which activities should have revenue raising as their goals. The educational materials, for instance, could be considered like the television programs and measured only in terms of educational impact and institutional innovation, but not in terms of revenue. Second, it is not clear which activities must also have some educational and innovative impact as well. The nonbroadcast products, for instance, might increase revenues by raising prices if relieved of the goal of maintaining artificially low prices to cater to low-income families. The rationale for the new policy might be to maximize revenues from nonbroadcast products so that the needs of low-income families might be better served by the broadcast programs.

The relative ease with which the financial impact can be measured also tends to conceal the problem of judging the soundness of a business venture. In addition to the amount of revenue, commercial ventures should also be judged by their market penetration and potential market. For the purpose of ensuring steady revenues for CTW, various ventures also have to be balanced between long- and short-term returns.²⁵

Summary of Goal-Oriented Impact Measures

As a brief review of the impact measures discussed thus far, Figure 3 shows the measures in relation to the three goals. Before applying such measures, CTW would have to agree on a set of goals and decide which of its activities has which goals or combination of goals. Then, each activity would be assessed by at least one measure reflecting the progress in achieving its goal(s).

MEASURING THE POTENTIAL NEGATIVE IMPACT OF CTW ACTIVITIES

The impact measures identified thus far tend to reflect positive achievements in line with CTW's goals. While CTW or any other organization would not intentionally seek to have a negative impact, most social activities can produce some negative impact. A comprehensive assessment of CTW must therefore attempt, at a minimum, to detect areas of negative impact. For CTW, such areas might include the

²⁵ Surprisingly, there appear to be few criteria for assessing the financial impact other than purely business criteria. The use of the venture capital from the Ford Foundation, for instance, could have included stipulations about the money being used for socially beneficial or innovative projects. Ford officials consulted by the author, however, all indicated that the sole criterion was the soundness of a project as a business venture, and that such a venture could include such innovation-poor projects as the purchase of land (e.g., for office and studio space), and the use of the capital as leverage for loans or for any number of commonplace business purposes.

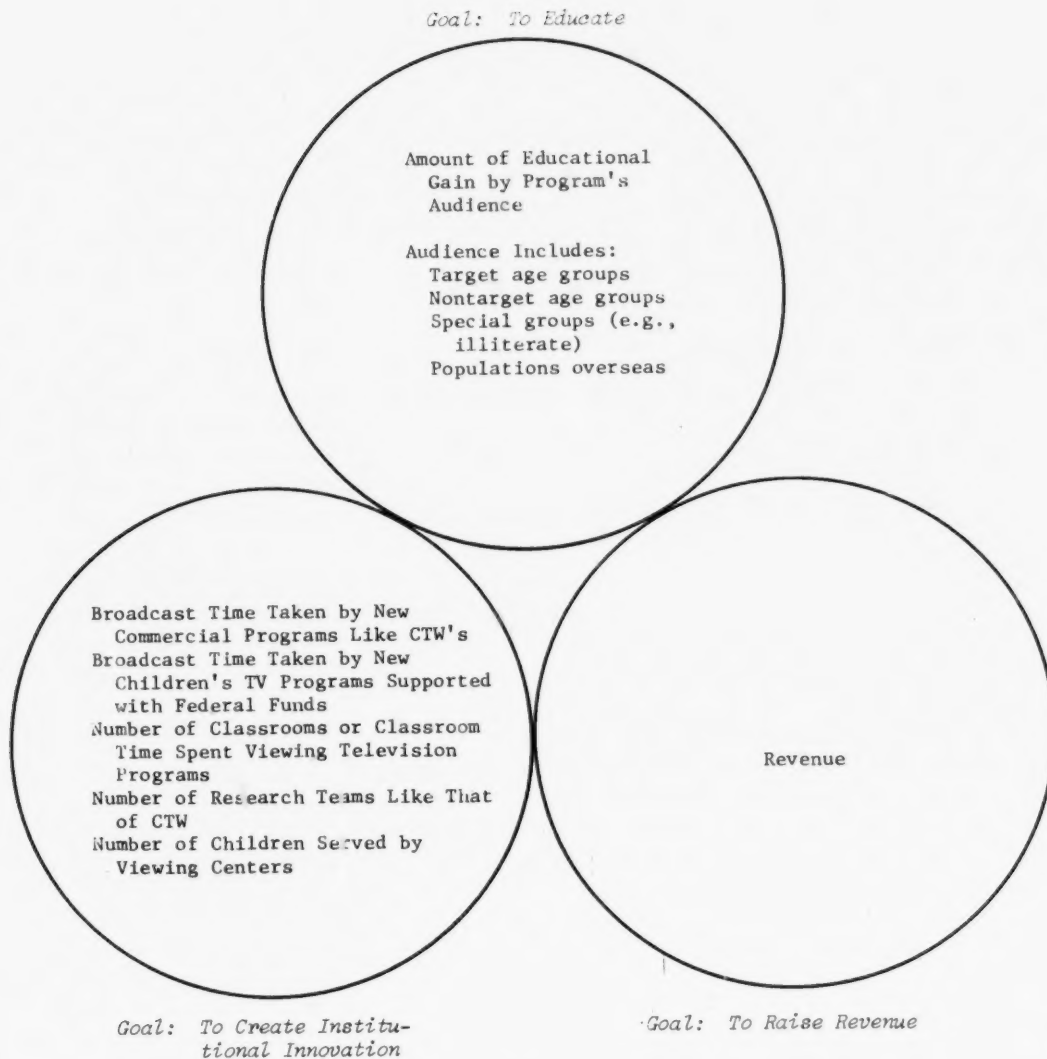


Fig. 3—Potential measures for assessing three goals

untoward effects of television viewing, of institutional innovation, and of cognitive changes.

On the matter of television viewing by children, frequent comments on *Sesame Street* and *The Electric Company* have been that both programs promote two undesirable by-products: a short attention span leading to boredom in school (due to the programs' magazine formats and emphasis on highly stimulating but short segments), and a propensity to watch more television in general or to give undue credence to what is shown on television. Though these effects may be readily produced by other television programs as well, British broadcasting officials were par-

ticularly sensitive to them in deciding whether to show *Sesame Street* in Great Britain. One American television critic has described the problem in the following manner:

When I watch *Sesame Street*, I find myself singing along . . . what a W is and all that . . . It has all the good teaching techniques and it also has all the techniques of commercials. But I'm also scared to death by *Sesame Street*. I think you can take 12 million kids and brain wash them in no time flat. It's quite all right when you're teaching them the alphabet . . . but you go on from there, to teach them certain values . . . this thing has been so researched, it gives one a terrible "Big Brother" feeling.³⁶

On the institutional side, CTW's relatively innovative approach in studying viewer behavior and conducting "distractor" tests could also have an ambivalent effect. This would occur if the networks used such techniques to produce attractive programs without regard to educational value or effects on attitudes or behavior.

Finally, even the achievement of cognitive gains also involves a potentially ambivalent impact. Persons concerned with educational inequality have suggested that *Sesame Street*, for instance, may have aggravated the gap between middle- and low-income educational achievement. This may have occurred because the program appears to produce greater cognitive gains among the children of middle-income families,³⁷ even though all children do derive some cognitive benefit. In terms of the low-income child's educational goals, then, *Sesame Street* may create an ambivalent effect.

Few of these potentially negative effects of CTW can be measured. Certainly much more research, to establish measures and to test the appropriate causal links, is needed before the importance of the negative effects can be understood.³⁸ The topic is worth exploration, however, and does raise an important element of doubt regarding the methodological limitations in assessing CTW's impact. This problem shall be pursued further in Section IV.

SUMMARY

This section has suggested that the most appropriate guide for measuring CTW's impact is to examine areas that reflect its potential organizational goals. Three general goals have been identified, with each of CTW's twelve activities tentatively placed in the context of one or more of three goals: to educate, to create institutional innovation, and to raise revenue. The identification of the goals leads

³⁶ Judith Crist, as quoted in Norman S. Morris, *Television's Child*, Little, Brown, Boston, Massachusetts, 1971, p. 167.

³⁷ Thomas Cook of Northwestern University is currently examining this issue. For further discussion of his work, see the next section of this report. One interesting point is that *Sesame Street* may have raised the country's entire preschool achievement norm, as measured by standard reading readiness tests. If this is so, such a change may be viewed positively (as further evidence of the truly broad educational impact of the program) or negatively (as evidence that low-income children have fallen farther behind their middle-class peers).

³⁸ One preliminary attempt to broaden the inquiry, though not necessarily focusing on negative effects, is Natalie Sproull, "Visual Attention, Modeling Behaviors, and Other Verbal and Nonverbal Metacommunication of Prekindergarten Children Viewing *Sesame Street*," *American Educational Research Journal*, Vol. 10, Spring 1973, pp. 101-114.

to the types of measures that should be used for CTW, and this section has discussed a potential array of such measures. The section has also raised the problem of measuring the potentially negative effects of CTW's activities.

Our attention now turns to the possible sources of data to be used in such measures, and to the adequacy of existing evaluations of CTW.

III. EXISTING DATA ON THE CHILDREN'S TELEVISION WORKSHOP

During the last few years, formative and summative evaluation have become an integral theme in the development of educational programs.³⁹ Federal programs to improve education have increased rapidly, and policymakers are constantly confronted with the problem of deciding which programs to initiate, which to continue, and what program variations to test. Evaluation serves as the primary attempt to provide "objective" feedback in this decisionmaking process, although many other political and subjective factors inevitably play an important role as well.

The Children's Television Workshop (CTW) already compares favorably with other organizations in the extent to which it has fostered the use of evaluation. CTW has commissioned several major evaluations itself, and has cooperated in evaluations sponsored by others. Most of these evaluations concentrate on the educational rather than innovative or revenue-raising achievements of CTW. Although the evaluations do not all reach the same conclusions, and although some have been criticized, the basic fact remains that CTW's activities have offered enough social innovation and documentation to sustain serious academic attention and debate. Prior to considering any new attempt to assess CTW, then, it is necessary to review briefly the existing evaluations, with the major questions being whether any of the existing studies suffice, and whether any of the data would be suitable for use in a new assessment of CTW. Because the review is concerned mainly with impact evaluations, the topic of formative research has been ignored, although in fact CTW has put a substantial effort into formative research.⁴⁰

The evaluations to be discussed below include: (1) an independent study carried out by Herbert Sprigle, (2) an independent study conducted by Herman Land, (3) audience surveys carried out by national polling organizations, (4) three separate studies done by the Educational Testing Service (ETS), and (5) a study by Thomas Cook that reanalyses the ETS data.⁴¹

³⁹ The terms "formative" and "summative" were coined in M. Scriven, "The Methodology of Evaluation," *AERA Monograph Series on Curriculum Evaluation*, No. 1, 1967, pp. 39-83. A formative evaluation refers to the collection of evidence during the development of a new curriculum in such a way that revisions of the curriculum can be based on this evidence. A summative evaluation is used at the end of an educational course or program to determine the effectiveness of that course or program.

⁴⁰ For example, see Barbara Frenzel Reeves and Edward L. Palmer, "The First Year of *Sesame Street*: The Formative Research," Children's Television Workshop, New York, December 1970.

⁴¹ This brief review does not cover in any detail the findings of the various evaluations. The reader should be aware that the ETS evaluations are substantial pieces of research, and have drawn equally substantial comments and scholarly attention.

SPRIGLE STUDY

Herbert Sprigle, the director of an early childhood in-school program known as "Learning to Learn," conducted an experiment in *Sesame Street*'s first year, comparing the effects of *Sesame Street* with the effects of an in-school kindergarten program. While an initial study consisted of a cross-sectional comparison between the two treatments, Sprigle subsequently conducted a follow-up study that allowed for longitudinal comparisons as well.

Sprigle's original study compared 24 matched pairs of disadvantaged children. One member of each pair was placed in a preschool program developed around viewing *Sesame Street*; the other member was placed in an alternative preschool program. Testing on the Metropolitan Readiness Test revealed that, after the treatment period, the *Sesame Street* group performed worse than its mates, being on the average more than 30 IQ points worse than the non-*Sesame Street* group. The second study compared children exposed to a *Sesame Street* program for two years with matched groups in Head Start programs, and yielded similar results.⁴² Sprigle has concluded from these studies that *Sesame Street* has little, if any, educational effect.

Unfortunately, the studies are seriously inadequate for drawing any conclusions. First, the very notion of considering *Sesame Street* as part of an in-school program, to be compared with other in-school programs, is an incorrect interpretation of the use of television as a supplementary form of education. The pertinent comparison would have been viewing versus nonviewing under otherwise identical situations. Second, the basis for the matched groups is not made clear; in particular, there is considerable suspicion that the groups differed substantially in their pre-test scores (these scores were not given in the report of the study), and that this difference was more likely to account for a subsequent 30-point IQ gap rather than any exposure to television. Third, the measurement error could easily have been large because of the small sample, the fact that the pre- and post-test instruments were different, and the use of five- and six-year-olds rather than the four-and-a-half-year-olds targeted by *Sesame Street*. In short, the results of the Sprigle study deserve no further attention unless the study is replicated under far improved experimental conditions.⁴³

LAND STUDY

The U.S. Office of Education commissioned Herman Land to conduct a broad review of CTW in 1971. His purpose was primarily to identify those characteristics of CTW that were most responsible for its success, and to determine how such characteristics could be reproduced elsewhere. The Office of Education wanted to know, in other words, how the CTW "model" could be replicated so that additional television programs could be created for educational purposes.

⁴² Herbert Sprigle, "Who Wants to Live on *Sesame Street*?" *Childhood Education*, Vol. 49, No. 3, December 1972.

⁴³ A brief critique of Sprigle's original study can be found in Gerry Ann Bogatz and Samuel Ball, *The Second Year of Sesame Street: A Continuing Evaluation*, Educational Testing Service, Princeton, New Jersey, 1971.

The Land study provides a comprehensive and helpful description of CTW's activities.⁴⁴ It covers both the broadcasting and nonbroadcasting aspects, even giving some attention to such purely administrative matters as financial management. The study has no analytic framework, however, and is based on the author's impressions and his bias, openly stated in the preface, in favor of CTW. Thus the study fulfills only a descriptive function, with a special emphasis on the early history of CTW. The lack of an analytic framework detracts from the impact of Land's conclusions: CTW's success is attributable to good timing (*Sesame Street* filled a void), talented and well-motivated people, and the availability of large sums of money.

AUDIENCE SURVEYS

Since the first year of *Sesame Street*, several surveys have attempted to determine the size and demographic nature of the viewing audience. These surveys have been carried out by Louis Harris and Associates for the Corporation for Public Broadcasting, and Daniel Yankelovich, Inc. for CTW, the latter including viewer statistics by the A.C. Nielsen Co. These surveys cover different time periods and different target populations. The Yankelovich surveys, for instance, were designed specifically to determine the amount of viewing by urban low-income audiences, and therefore focused on four inner-city neighborhoods: two in New York City, and one each in Chicago and Washington, D.C.

Each of these survey organizations uses different sampling techniques and measurement instruments that make generalization of the results to a national population difficult. The Nielsen ratings, for example, are based on actual viewing records from an attachment to the television set. However, the rejection rate among potential respondents and hence the true nature of the Nielsen sample are unknown. In general, the available reports from these surveys do not describe the methodological procedures or present the findings in sufficient detail. The unknown quality of the samples has meant that the *national* characteristics of CTW's viewing audience are really unknown. Moreover, each of the existing surveys only assessed viewing at a single point in time; the surveys did not inquire about the amount of viewing over several seasons, to determine how many of *Sesame Street*'s programs are watched by the average viewer.

The major purpose of an audience survey is to determine how many people are watching CTW's programs. The answers are based in part by the nature of the sample, the definitions used for "watching," and other methodological factors, e.g., the season in which the survey is undertaken. (The sampling problem, incidentally, is further aggravated by the erratic availability of public television stations.) The results of the surveys suggest that: (1) the audiences of *Sesame Street* and *The Electric Company* have been increasing and approach the audience size of popular commercial programs, (2) *Sesame Street* is probably watched by a majority of the nation's children ranging from two to five years of age, and (3) *Sesame Street* is probably watched by a majority of the two- to five-year-old children of low-income families. More precise conclusions are not possible from the existing survey data.

⁴⁴ Herman W. Land, *The Children's Television Workshop: How and Why It Works*, Nassau Broad of Cooperative Educational Services, Jericho, New York, 1972.

Audience surveys will play an important role in any future assessment of CTW's impact. No matter how the educational impact is measured, accurate estimates of the number of viewers, including the duration and intensity of viewing, will be needed. In addition, a properly designed survey, whose sample characteristics are amenable to valid generalizations to a national population, can also provide information on the number and the nature of nontarget population viewers. Unfortunately, it is likely that there are no further benefits to be derived from reanalyzing the existing survey data, and new audience surveys will have to be designed and conducted.

We turn now to three evaluations that focus entirely on the cognitive issue, i.e., the effects of CTW's television programs on learning. These three studies were carried out by the Educational Testing Service (ETS), and sponsored by CTW.

FIRST ETS EVALUATION OF SESAME STREET

The first ETS evaluation tested 943 children, aged three to five years and drawn heavily from disadvantaged backgrounds, at five different sites: Boston, Massachusetts; Phoenix, Arizona; Durham, North Carolina; Philadelphia, Pennsylvania; and a rural area in northeastern California.⁴⁵ A specially designed test was administered to these children before and after six months of viewing *Sesame Street* during its first season, 1969-1970. The test consisted of 203 questions aimed at eight curriculum topics: body parts, letters, forms, numbers, relational terms, sorting, classification, and puzzles. All eight were explicit parts of the first year's *Sesame Street* programs, and thus the test constituted a criterion-referenced test to assess the specific goals of the program rather than general cognitive achievement.

Because of initial fears that not enough children would view a new program such as *Sesame Street*, the experimental design for the evaluation randomly assigned the children at each site into two groups of at-home viewers: those explicitly encouraged to view the program through publicity and field visits, and those not encouraged. In actual fact, the original fears were completely unfounded. Many more children than expected, in both the encouraged and nonencouraged groups, viewed the program; if anything, ETS had difficulty identifying a suitable group of children who had actually never or rarely seen the program. The experimental design also included a similar division of children in a school situation, with encouraged classes receiving television sets and teacher guidance, and with other classes receiving no encouragement. Assessment of the ultimate amount of viewing for each child involved a test of knowledge about the program's characters, a post-test adult survey, and periodic spot checks.

The results of this evaluation, as reported by ETS, were the following: the amount of pre- to post-test gain increased in relation to the amount of time children watched the program; the gains were as great if not greater at home than in school, so that teacher supervision appeared unnecessary in producing the results; and the

⁴⁵ Samuel Ball and Gerry Ann Bigata, *The First Year of Sesame Street: An Evaluation*, Educational Testing Service, Princeton, New Jersey, 1970.

gains were greatest among the three-year-olds and least among the five-year-olds. The results received considerable publicity, with the overall conclusion being that learning increased with exposure to *Sesame Street*.

Unfortunately, the evaluation design and measurement instruments had sufficient flaws to raise questions about the significance of these conclusions. First, the samples at each site represented different socio-economic characteristics and were of considerably different size (Boston had over 400 children, and California less than 100). The extent to which the findings at any of these sites can be generalized to larger populations is unclear. Second, the measures used to assess viewing time were not adequate. Testing for knowledge about the program's characters, for instance, clearly confounds viewing with learning, and does not provide an accurate assessment of viewing time.

Third, the "treatment" consisted not of viewing, but of encouragement to view. This meant that while children were randomly assigned to the encouragement/nonencouragement groups, they were not so assigned in terms of viewing time. This is extremely important, as it turned out that the amount of viewing time was directly (and significantly) correlated with pre-test scores. The simple result that learning increased with viewing, then, could have been produced by both the contaminating effects of "encouragement to view" and the problem that the children who viewed more, by self-selection, may have already been intellectually different from the children who viewed less. Supplementary analyses, holding prior achievement levels constant, showed that the effects of viewing were still statistically significant, but the ETS report readily admits that such analyses do not provide a definitive answer.⁴⁶

ETS attempted to address some of these problems in the introduction to its second study.⁴⁷ In addition, Thomas Cook is conducting a systematic reanalysis of the ETS data, holding "encouragement to view" and other contaminating factors constant, and then attempting to discern the effects of viewing on learning.⁴⁸ This is an important procedure since most children who view *Sesame Street* will not be part of any experiment in which they are exposed to an "encouraged" treatment. (At the same time, no one has actually surveyed the amount of "encouragement" in a normal, nonexperimental home.)

In summary, the first ETS evaluation does not, in its originally reported form, provide a complete test of the impact of viewing *Sesame Street*. While this discussion has focused primarily upon the weaker points in the evaluation design, it should also be pointed out that the design may also have produced an underestimate of *Sesame Street*'s effects, in that viewing time was limited to six months, and most children may watch *Sesame Street* for two or more seasons. In retrospect, the evaluation design cannot be overly criticized, what with the normal difficulties of large-scale field research compounded by the fact that many more people watched *Sesame Street* in its first year than anyone, including CTW, would ever have expected.

⁴⁶ Ball and Bogatz, *The First Year of Sesame Street*, p. 368.

⁴⁷ Gerry Ann Bogatz and Samuel Ball, *The Second Year of Sesame Street: A Continuing Evaluation*, Educational Testing Service, Princeton, New Jersey, 1971.

⁴⁸ Cook kindly made all the information cited about his study available through personal communication with the author of the present document.

SECOND ETS EVALUATION OF SESAME STREET

The second ETS evaluation of *Sesame Street* took place the following year (1970-1971), and consisted of two separate parts. The first part was an attempt to replicate the first year's study, by examining the effects of viewing and nonviewing on disadvantaged children at two new sites, Winston-Salem, North Carolina, and Los Angeles, California. The second part was a follow-up of the Boston, Durham, and Phoenix children studied in the first year. However, the analysis of the second part, while yielding some results, was confounded both by the inadequacies of the first year's evaluation design and by the fact that virtually all the sampled children (encouraged and nonencouraged) were viewing *Sesame Street* by this second year. Hence the follow-up part of the study will not be further discussed.

The selection of the two new sites was determined by the fact that *Sesame Street* in its second year of broadcasting was not readily available to many homes in these two cities. In Winston-Salem, the show was not broadcast at all, and the evaluation was conducted by using the facilities of the local cable television operator (homes were randomly assigned to viewing and nonviewing conditions). In Los Angeles, the show was only available over an ultra high frequency (UHF) channel, and the experiment consisted of identifying homes not receiving this channel, distributing the correct tuners to a sample of these homes, and observing a control group of homes with no such tuners. In this second study, the "treatment" still included encouragement to view *Sesame Street*; this time, however, the treatment was more effective, as only 9 of 130 encouraged children did not watch *Sesame Street*, and 54 of 153 nonencouraged children occasionally viewed the program.

The results, on a battery of criterion-referenced tests similar to those of the first year, showed that the cognitive gains from pre- to post-test were significantly greater among the encouraged group. The differences on some of the subscales of the test were greater than on other subscales; however, in all cases the differences were in the hypothesized direction. Interestingly, an analysis separating the effects of viewing from the effects of encouragement showed that both had (statistically) independent and significant impacts on cognitive gains. The major conclusion again was that learning had increased as a result of viewing *Sesame Street*.

This second ETS study did not have the same design problems as the first. Comparisons of pre-test scores, for instance, showed that the viewing and nonviewing groups did not differ significantly before the experiment began. On the other hand, this second study covered only 283 children at two sites, so that the basic generalizability of the results may be questioned. Moreover, the measurements of the amount of viewing, while an improvement over those used in the first year, were still based primarily on verbal reports and recall of the previous day's television viewing. In general, however, the second ETS study provides considerable, if not definitive, evidence that children learn from *Sesame Street*.⁴⁹

⁴⁹ Gerry Ann Bogatz has provided a self-analytic description of the conduct of the first two ETS evaluations. See "Evaluations of *Sesame Street*: Two Case Studies," Educational Testing Service, Princeton, New Jersey, manuscript in preparation.

THIRD ETS STUDY: EVALUATION OF THE ELECTRIC COMPANY

The first year's broadcasting of *The Electric Company* (1971-1972) provided the opportunity for ETS's third study. The study examined the effects of viewing *The Electric Company* on both "in-school" and "at-home" situations, with two pairs of sites chosen for both of these situations. For the in-school test, the sites were Fresno, California, and Youngstown, Ohio. The selection of sites was limited to those cities where *The Electric Company* could not be seen at home, and Fresno and Youngstown were chosen to cover two different types of populations: Fresno was considered more rural and included a Spanish background population in one-half of the sample, and Youngstown was considered more urban and included a black population in one-half of its sample. For the at-home test, the sites were Richmond, Virginia, and Washington, D.C., with the latter including a large proportion of black children.

The testing was planned to include 100 classrooms of children in grades 1 through 4 at each site. The sample is shown in Table 6. At each site, classrooms within each grade were randomly assigned to an experimental and a control condition. For the in-school test, the experimental condition consisted of daily classroom viewing of *The Electric Company* for six months, accompanied by the use of a teacher guide and classroom discussion; the control condition consisted of no program. For the at-home test, the experimental condition called for teacher encouragement for the children to view *The Electric Company* at home; the control condition consisted of no mention of the program by the teachers.

A wide variety of tests was administered to all children before and after "treatment," including a specially designed criterion-referenced test of 123 items (with 19 subscales) administered as a group test, an individual test of 42 items administered to a subsample of children, the Metropolitan Achievement Test, and attitudinal questions. The criterion test focused on reading skills: blending of letter sounds, letter grouping, scanning for structure, and reading for meaning. In addition, separate questionnaires were administered to parents and teachers, records were kept of school attendance and classroom activities, and separate viewing records were kept by children and parents in the at-home test.

The evaluation produced the following results. First, the at-home treatment was not properly established, as a high proportion of the children in the nonen-

Table 6
SAMPLE IN ETS EVALUATION OF THE ELECTRIC COMPANY

Site	Number of Classrooms	Number of Children
In-school viewing		
Fresno, California	92	1,887
Youngstown, Ohio	100	2,541
At-home viewing		
Richmond, Virginia	98	2,287
Washington, D.C.	86	1,648

couraged or control condition unexpectedly viewed the program at home. In retrospect, the design of the at-home test may have been rather naive, since it could have been expected that many children would normally be inclined to watch *The Electric Company* as a result of *Sesame Street*'s success, and that the children in the experimental and control conditions would communicate outside the classroom and hence undermine any treatment condition based merely on a teacher's verbal suggestion.

Second, the in-school situation yielded a wide variety of results, only the most important of which will be highlighted. The in-school situation showed that children exposed to *The Electric Company* improved more, especially in the early grades, on the criterion test than children in the control condition. The significance of this result varied for each subscale, but in all cases the result was in the hypothesized direction. On the Metropolitan Achievement Test, only in the first grade at Youngstown and among a special subgroup of poor readers in the fourth grade at Fresno⁵⁰ did viewers improve significantly more than nonviewers. No important differences between experimentals and controls were found on the other measures (school attendance, parent questionnaire, teacher questionnaire, and classroom observations). Once again, the major conclusion was that the television program had the desired effect in increasing children's learning abilities.

The results of this third ETS study, however, are not really so readily interpretable. Although over 4,000 children were tested in school, the evaluation only covered two sites. More important, missing from the study is a valid control group, i.e., one with a sham television program where concomitant sham teacher guides were distributed. Given the practical difficulties of experimentation in a school system, this lack of such a control group is perhaps forgivable; however, the study as designed leaves open the possibility that the cognitive gains could be attributable to other classroom conditions created by the fact of viewing, and not merely the exposure to the program itself.⁵¹ Additional problems may arise in interpreting the results because of other methodological reasons: e.g., the unit of analysis was the classroom, and not the individual student; and the pre-test scores of the third and fourth grade students on the criterion test were already very high (the same criterion test had been given to students in all four grades).

Needless to say, the final judgment on the effects of *The Electric Company* is still forthcoming. A fourth ETS evaluation, examining the second year of *The Electric Company*, may not shed much additional light on this question. The new study, to be reported during the spring of 1974, only includes a follow-up of the third study, and does not extend the evaluation to new sites or treatment groups.⁵²

⁵⁰ The analysis of the results consistently examined separately the effects of viewing on the real target population, i.e., students defined as all 1st graders, together with those scoring in the lower half of the Metropolitan Achievement pre-test in the second grade, and those scoring in the lowest quartile of that pre-test in the third and fourth grades.

⁵¹ This shortcoming may not be as important in Fresno, where classrooms normally used television, and where the control group therefore watched programs other than *The Electric Company*, but not necessarily with the same role for teachers' guides and class discussion. The Youngstown sample, however, had no television in the control group.

⁵² Gerry Ann Bogatz of ETS kindly described the fourth year study design via personal communication with the author.

COOK'S REANALYSIS OF THE ETS DATA ⁵³

As previously mentioned, an independent reanalysis of the ETS data is being carried out by Thomas Cook of Northwestern University. Cook's reanalysis thus constitutes another type of evaluation of CTW's activities, with the focus on *Sesame Street* and the first two ETS studies.

Cook's major concerns are methodological and public policy-related. As for methods, Cook's study will hopefully offer a useful critique of ETS's measurement devices and their shortcomings. His study will also examine the experimental design used in the ETS studies, and especially the degree to which the effects of encouragement and viewing were confounded.

As for public policy, Cook's overriding preoccupation is with the inequality of education issue, and the fact that, to the degree *Sesame Street* is effective, it may widen the achievement gap between socially advantaged and disadvantaged children. Cook's analysis therefore will also concentrate on the amount of cognitive gain (whether produced by encouragement, by viewing, or both) among disadvantaged as opposed to advantaged children, with the hypothesis that viewing *Sesame Street* may aggravate these differences (though both groups of children show gains). Whatever the outcome of Cook's study, it should be noted that CTW has made few claims on behalf of specific subgroups of children; its major goal remains the raising of cognitive achievement of all children, with no specific claims of reducing the educational inequality gap.⁵⁴

Cook's policy-related concerns may also extend beyond the ETS data. For instance, he is also interested in commenting on the other activities of CTW besides *Sesame Street*, such as CTW's cable television venture.⁵⁵ Any judgment on the nature of Cook's work must obviously await the completion of his study.

SUMMARY OF EXISTING EVALUATIONS

This brief review of existing evaluations suggests that none is sufficient, on its own, to provide an assessment of CTW's full impact. The single study that is sufficiently comprehensive (the Land study) actually does not provide any measurement of CTW's activities. The audience surveys by national polling organizations provide a critical piece of evidence, but need to be repeated under more rigorous methodological conditions so that precise estimates can be made of the total viewing audience and its characteristics. The exhaustive ETS evaluations, while providing ample

⁵³ Again, the information about this study was obtained through personal communication.

⁵⁴ While CTW itself has made few such claims, the situation was not clarified by a recent article by the ETS evaluators. The article begins by asserting that *Sesame Street* might serve the goals of compensatory education; only after the results are reported is compensatory education redefined to include middle-class as well as disadvantaged groups. See Samuel Ball and Gerry Ann Bogatz, "Research on *Sesame Street*: Some Implications for Compensatory Education," in *Proceedings of the Second Annual Hyman Blumberg Symposium on Research in Early Childhood Education*, Johns Hopkins Press, Baltimore, Maryland, 1972.

⁵⁵ For recent discussions of urban cable systems and their potential service applications, see Walter S. Baer, *Cable Television: A Handbook for Decisionmaking*, The Rand Corporation, R-1133-NSF, February 1973; and Robert K. Yin, *Cable Television: Applications for Municipal Services*, The Rand Corporation, R-1140-NSF, May 1973.

measures, are too narrow in that they examine only the target population's cognitive gains resulting from the television programs. Moreover, the ETS data are limited to a few sites, from which any generalizations about national populations cannot be accurately made.

A totally new evaluation would thus be needed for at least two reasons. First, the design could not only cover CTW's impact on the target population of viewers, but also its impact on nontarget viewers (e.g., the effect of *The Electric Company* on four-year-olds, or the effect of either television program on slow-learning children), on television broadcasting more broadly, and on educational institutions more broadly. Second, the evaluation could address several of the methodological issues raised by the ETS and Cook studies, and attempt to settle the existing controversies. The hope that a new evaluation will produce a definitive outcome, however, should be tempered by the generally crude state of the art in evaluation methodology. Even if the existing methodological issues were addressed properly, any new evaluation would inevitably raise new questions.

For instance, the research on television violence and its effect on aggression in children is already well developed, yet has still been unable to answer the basic question. Much of the research carried out as part of the Surgeon General's study on television and social behavior pointed to a causal link between violence viewing and aggression. Yet the evidence was not so strong as to be convincing. As one result, the Advisory Committee to the Surgeon General (which included members of the television industry) concluded that the studies suggested:

... a preliminary and tentative indication of a causal relationship between viewing violence on television and aggressive behavior; an indication that any such causal relationship operates only on some children (who are predisposed to be aggressive); and an indication that it operates only in some environmental contexts.⁵⁶

Most recently, several investigators have even called into question the use in these studies of the standard experimental or survey approaches, advocating the need for multi-wave panel studies (i.e., making several pre-tests and several post-tests) in relation to any exposure to television programs.⁵⁷

The point here, however, is not merely to single out the research on violence and aggression. Rather, the issue is one of general evaluation methods, and the ensuing consensus (or lack of it) over evaluation results. Major inquiries in the past, whether focusing on television and aggression, smoking and health, the effects of pornography, or equal educational opportunities, have simply not led to the clear-cut answers expected by policymakers. Rather, these inquiries have tended to reveal cleavages within the academic community, have drawn sometimes unjustified political fire, and have produced a continuing array of studies rather than convergence over a single set of answers.⁵⁸

⁵⁶ Report to the Surgeon General, *Television and Growing Up: The Impact of Televised Violence*, U.S. Government Printing Office, Washington, D.C., 1972, p. 11.

⁵⁷ J. Ronald Milavsky, et al., "Exposure to TV Violence and Aggressive Behavior in Boys, Examined as Process: A Status Report of a Longitudinal Study," paper presented at the American Sociological Association annual meeting, New Orleans, August 1972.

⁵⁸ Some of the controversial nature of the Surgeon General's study is reported in Leo Bogart, "Warning: The Surgeon General Has Determined That TV Violence is Moderately Dangerous to Your Child's Mental Health," *Public Opinion Quarterly*, Vol. 36, Winter 1972-1973, pp. 491-521.

POTENTIAL PROBLEMS FOR A NEW EVALUATION

Several other important problems would also confront a new evaluation. In the end, these would make a new evaluation, for the purposes of assessing CTW's overall impact, extremely difficult.

First, while a survey can and should be conducted to determine the characteristics of the national viewing audience, it will not be easy to identify a national population sample for the purpose of designing new experiments on the *effects* of viewing. There is little chance of identifying an appropriate sample of children who have not seen *Sesame Street* or *The Electric Company* and who could therefore serve in a control group. Nonviewing populations, whether in this country or even in Canada, are likely to be anomalous and hence unrepresentative of the national population, simply because the programs are now viewed so widely. The best that a new evaluation could do would be to focus on a few more sites like Winston-Salem and Los Angeles, and report the results of viewing and nonviewing for those sites. But such results still could not be extrapolated to a national population.

Second, even if adequate nonviewing samples are found, the existing evaluations suggest that it will be difficult to make the experimental manipulation, i.e., ensure that the viewers and nonviewers remain in the intended experimental and control groups. Truly effective manipulation, in other words, has yet to be carried out, and the cost of a new evaluation would have to be weighed against the risks of an abortive design.

Third, the most satisfactory measure of actual viewing time is not any kind of verbal report, but actual recordings of frequency of viewing. The Nielsen ratings are based on such recordings, but the methods are expensive and still do not reflect the quality of the viewing that has occurred. (In a strict sense, what Nielsen measures is *television* behavior, i.e., whether the television is on or off, and not *human* behavior, i.e., the duration and quality of viewing.) Thus almost any new evaluation will likely be susceptible to challenge on the grounds of the measures used for viewing time.

Fourth, new surveys would have to be carried out to estimate the size and characteristics of the viewing audience on a national basis.

Finally, new research would have to be carried out so that, whatever the cognitive results, they could be given a social value, such as the previously described equivalent in-school costs. At a minimum, this new research would involve some field testing and thus an additional cost factor.

IV. CONSIDERING A NEW ASSESSMENT OF THE CHILDREN'S TELEVISION WORKSHOP

The previous sections of this paper have described the breadth of CTW's activities, suggested potential measures of CTW's impact, and reviewed the adequacy of some existing data. The analysis has thus far uncovered a few important methodological problems. The analysis has not addressed, however, certain fundamental issues that transcend these problems, and that must now be faced in considering any new assessment of CTW.

Such an assessment inevitably raises two major questions:

- What is the purpose of a new evaluation, and who is the audience?
- Can the impact of a single CTW activity, much less the impact of all of its activities, be aggregated in some benefit/cost or similar summary fashion?

The purpose of this section is to suggest the problems in answering these questions, and to show how such answers will very closely determine the ultimate nature of any assessment of CTW.

POSSIBLE PURPOSES FOR AN EVALUATION AND ITS POTENTIAL AUDIENCES

This first issue is the key to any further work on evaluation design.⁵⁹ The purposes and audience for an evaluation determine the questions to be raised by the evaluation, the amount of money that is justifiable for the evaluation, and the degree of detail and accuracy required. As one example, the costs of an evaluation are directly related to the size of the sample to be studied, and the size of the sample is inversely related to the measurement error in the data. Therefore, the justifiable costs for an evaluation depend on the types of decisions to be made as a result of the evaluation, and by the accuracy of the information needed to make these decisions.

⁵⁹ The author is grateful to Joseph S. Wholey of The Urban Institute for his brief but incisive personal comments on evaluation issues. For another useful discussion of evaluation issues, see R. A. Levine and A. P. Williams, Jr., *Making Evaluation Effective: A Guide*, The Rand Corporation, R-788-HEW/CMU, May 1971.

Joseph Wholey of The Urban Institute has defined four possible types of evaluation that serve four somewhat different decisionmaking needs:⁶⁰

- Program impact evaluation, which assesses the overall effectiveness of a program;
- Program strategy evaluation, which assesses the relative effectiveness of different strategies used within a program;
- Project evaluation, which assesses the effectiveness of an individual project in achieving its stated objectives; and
- Project rating, which assesses the relative effectiveness of different local projects.

This typology was designed primarily for broad federal programs, such as the antipoverty program, and not necessarily meant for single organizations such as CTW. At the same time, however, the typology captures the critical questions that need to be answered in any evaluation of CTW. Let us therefore examine three different types of evaluation that draw from this typology and that might be applicable to CTW: an *impact* evaluation, a *strategy* evaluation, and a *project* evaluation.

Impact Evaluation

An impact evaluation of CTW would have as its main audience national decisionmakers who are concerned with education and the use of television for instructional purposes. Included among such decisionmakers might be the U.S. Congress, the U.S. Office of Education, the Corporation for Public Broadcasting, the Public Broadcasting Service, and foundation officials and other private financial sponsors of educational television programs. An evaluation serving these audiences would have to be directed at their resource allocation functions. It would provide information, in other words, to help them choose among different programs for financial support. Thus the major features of the evaluation would include not only an assessment of the impact of CTW, but also (1) a determination of the costs involved in supporting CTW activities, and (2) a comparative framework so that CTW's impact and costs could be judged relative to those of some other organization.

The Land study provides a good example of the breadth of an impact evaluation. In fact, the Office of Education sponsored the Land study because it was interested in developing new organizations, tailored after the so-called CTW model, to produce new educational television programs. To summarize, the impact evaluation assumes that:

- The audience for an evaluation consists of national sponsors of educational television programs;
- The purpose of the evaluation is to serve decisionmaking needs for allocating resources among major programs;
- The evaluation consists of an overall assessment of the costs and impact of CTW's activities; and
- The evaluation includes a comparison between CTW and alternative organizations competing for the same or additional funds.

⁶⁰ Joseph S. Wholey, et al., *Federal Evaluation Policy: Analyzing the Effects of Public Programs*, The Urban Institute, Washington, D.C., 1970.

Strategy Evaluation

A strategy evaluation would have as its main audience the corporate officials of CTW. The evaluation would attempt to serve the decisionmaking needs within CTW, providing guidance on those activities to encourage or discourage. For instance, if three activities (the cable television, television cassette, and feature-length film) all have the same goal of raising revenue, then a program strategy evaluation could compare these activities in order to identify the most effective means of achieving the goal. Similarly, other CTW activities could be assessed, and the information, if timely enough, could help in the formulation of CTW's periodic program plans.

The strategy evaluation would not, however, provide an overall assessment of CTW's impact. Nor would it provide any comparative data between CTW and other organizations. As a result of the narrower focus of this type of evaluation, CTW itself would probably have to sponsor the evaluation. Such self-support would help to ensure the relevance of the evaluation to serious decisionmaking needs, but would also probably add a bias in favor of presenting all of CTW's activities in the best light possible.

To summarize, the strategy evaluation assumes that:

- The audience for the evaluation consists of CTW's own decisionmakers;
- The purpose of the evaluation is to serve CTW's decisionmaking needs for developing and supporting its own activities;
- The evaluation does not necessarily include an overall assessment of the impact of CTW's activities; and
- The evaluation does not include any systematic comparison with other organizations or programs.

Project Evaluation

A project evaluation, like the impact evaluation, would again have national decisionmakers as its primary audience, and again serve decisionmaking needs with regard to allocating resources. However, unlike the impact evaluation, the project evaluation would focus on a specific CTW activity, and compare this activity with similar activities outside CTW. The project evaluation, in other words, would most eminently suit the current needs of the Office of Education or the Corporation for Public Broadcasting for evaluating and comparing specific project grants. Either of these agencies would thus be a likely candidate for sponsoring a project evaluation.

As an example, the Office of Education might support a comparative evaluation of the three popular preschool television programs: *Captain Kangaroo*,⁶¹ *Misterogers' Neighborhood*, and *Sesame Street*. The evaluation would compare the relative impact and costs of these programs. Although each program has different educational goals, the achievements of each could still be assessed, and the audience penetration and costs⁶² of each could also be measured. While the evaluation might not

⁶¹ Although *Captain Kangaroo* is a commercial program, educational segments have recently been produced and inserted with financial support from the U.S. Office of Education and the U.S. Office of Child Development.

⁶² Assessing the costs of a television program is difficult but can be accomplished. First, adequate data must be made available. Second, care must be taken to include comparable activities for each program. For instance, statements of the costs of producing *Sesame Street* usually include all research and overhead

avert the need to make important value judgments, it would serve to identify explicitly many judgments that are in fact now made implicitly.

The results of the evaluation would provide additional information for decisions on the future level of funding for these projects. The national sponsors would still have to make major value judgments concerning the desire to support, say, social-versus cognitive-oriented programs, but the evaluation might yield information about the relative effectiveness of the programs. The evaluation would not provide information on the overall impact of CTW's activities, or the activities of any other organization. However, the same evaluation framework could subsequently be applied to other programs, such as *The Electric Company* or the new health show. To summarize, a project evaluation assumes that:

- The audience for the evaluation consists of national sponsors of specific educational television projects;
- The purpose of the evaluation is to serve decisionmaking needs for allocating resources among major programs;
- The evaluation does not include an overall assessment of the costs and impact of CTW's activities; but
- The evaluation includes a comparison between specific CTW programs, such as *Sesame Street*, and other educational television programs.

Implications of Alternative Evaluations

These three types of evaluations alone present some important choices and trade-offs. Only the first type of evaluation will yield an overall assessment of CTW, but the costs of such an evaluation will be high and it will be difficult to decide what other organizations are good candidates for comparison with CTW. An obvious commercial choice would be the Walt Disney organization, but alternative nonprofit organizations are not easy to identify. The second type of evaluation would best address the management needs of CTW, but would not provide comparisons with other programs or even a full assessment of CTW's activities. This evaluation is likely to involve the lowest costs, however. The third type of evaluation would provide comparisons of specific CTW programs, probably at a cost somewhere between the first two types of evaluation.⁶³ Table 7 summarizes the comparative features of the three evaluations.

A major decision on the part of CTW and its sponsors, then, will concern the audience and scope of the evaluation. These factors, in turn, will determine the fundamental nature of the evaluation design. No matter what the final choice, however, all of the different types of evaluations do share one common assumption

items; statements of the costs of commercial programs such as *Captain Kangaroo*, however, exclude such items and other relevant network expenditures (see Alan Pearce, "The Economics of Network Children's Television Programming," mimeographed paper, Federal Communications Commission July 1972, pp. 10-12). Third, the measurement of costs should be based on the monetary total, and not take into account audience size, which confounds actual costs with potential impact and creates entirely different outcomes (see Rudy Bretz, *Three Models for Home-based Instructional Systems Using Television*, The Rand Corporation, R-1089-USOE/MF, October 1972, pp. 41-44).

⁶³ As a rough guideline, the two ETS evaluations of *Sesame Street* cost about \$500,000 in total. A new project evaluation, even using some of the existing data, is likely to cost at least the amount of one of ETS's evaluations, or \$250,000. The impact evaluation would thus be larger than this amount, and the strategy evaluation somewhat smaller.

Table 7
COMPARATIVE FEATURES OF THREE TYPES OF EVALUATION

Type of Evaluation	Feature of Evaluation				
	Audience and Sponsor	Purpose	Likely Costs ^a	Overall Assessment of CTW Included	Comparison with Non-CTW Activities Included
<i>Impact</i> evaluation To assess overall impact of CTW	National sponsors of children's television programs ^b	Serve national resource allocation decisions	High	Yes	Yes
<i>Strategy</i> evaluation To assess the relative effectiveness of strategies within CTW	CTW	Serve CTW decisions	Low	No, but possible	No
<i>Project</i> evaluation To assess the effectiveness of individual CTW activities	National sponsors of children's television programs ^b	Serve national resource allocation decisions	Moderate	No	Yes

^aFor a rough guideline on costs, see footnote 63.

^bFor example, the Office of Education, the Corporation for Public Broadcasting, or one of several private foundations.

that needs to be examined. This assumption is that the net impact of CTW's activities can, in some manner, be aggregated.

PROBLEMS IN AGGREGATING THE IMPACT OF CTW'S ACTIVITIES

Any comprehensive evaluation would require that a single CTW activity, such as the production of *Sesame Street*, be assessed by measuring all of its various impacts, and then in some manner would have these impacts added together to produce a summary measure of impact. An evaluation of all of CTW's activities further requires that this summary measure of each individual activity then be aggregated into a single, overall summary that reflects CTW's overall impact.

This notion underlies any consideration of the use, for instance, of benefit/cost analysis, whereby the costs of an activity are compared to its overall benefits. Thus a *hypothetical* assessment of social welfare programs might follow the matrix found in Table 8. The matrix illustrates how a series of poverty programs might be arrayed in order to reach conclusions about (1) the relative effectiveness of each program's penetration or coverage (line C of Table 8); (2) the costs and benefits of each program (lines D and G); and (3) ultimately the benefit/cost ratio of each program (line H). At a minimum, each program should have a benefit/cost ratio that has a value greater than unity, if the benefits and costs are measured in the same units.

Table 8

HYPOTHETICAL BENEFIT-COST MATRIX FOR POVERTY PROGRAMS

Anti-Poverty Programs											
Evaluation Data	Job Corps	VISTA	Community Action Programs						Delegated Programs	Non-OEO Project	
			Organization-Coordination Functions	Service Programs							
				Legal Services	Head Start	Follow Through	Upward Bound	Health Services			Others
A. Universe of need											
B. Program reach											
C. Program coverage (A/B)											
D. Program cost											
E. Cost per person reached (D/B)											
F. Cost of total coverage (E) x (A)											
G. Measures of program effectiveness											
1. Immediate objectives											
2. Poverty reduction											
H. Benefit/cost ratio (G/D)											

SOURCE: Joseph Wholey, et al., *Federal Evaluation Policy: Analyzing the Effects of Public Programs*, The Urban Institute, Washington, D.C., 1970, p. 37.

Thus for *Sesame Street*, the following effects would, *in theory*, have to be aggregated:

- Cognitive effects among the target population;
- Cognitive effects among nontarget populations;
- Innovation effects (e.g., number of new network programs that follow the same format or goals; incremental use of television in classrooms due to *Sesame Street* innovation);
- Revenue produced by foreign sales or production of the program;
- Possible effects of decreased attention span and possible detrimental effects of subsequent overinfluence by television; and
- Possible effects due to aggravation of the educational gap between children of low- and middle-income families.

The main conceptual problem with such an aggregation is that the various effects are measured in different units. Value judgments are required to compare cognitive gains with gains in institutional innovations.⁴⁴ Second, the effects may have a different impact on different populations. Finally, the problem of aggregation becomes even more difficult when an attempt is made to combine effects that are both positive and negative.

In short, any aggregate statement of the impact of a single CTW activity would be an extremely risky venture. Any suggestion of the impact of several CTW activities would naturally be even more hazardous.⁴⁵ As a consequence of this major conceptual difficulty, the previously identified methodological problems loom even larger. Those problems, to repeat briefly, involved (1) shortcomings with the existing ETS data so that only data from two sites, Winston-Salem and Los Angeles, may be worth reusing; (2) the almost impossible task of finding a new national sample of nonviewers of *Sesame Street*; (3) the difficulty of assuring an effective "treatment" for creating experimental and control groups; (4) the difficulty of assessing the amount of viewing without expending large sums of money; and (5) the likelihood that, because of the current state of evaluation research, any summative evaluation of such grandiose proportions is likely to stimulate rather than settle major academic controversies.

On the contrary, the success of benefit/cost or any other aggregate analysis is probably heavily dependent on a program having a single, unitary objective. Under such conditions, the analysis can help to choose among alternative strategies by assessing their effectiveness solely with regard to that objective. In the case of a program such as *Sesame Street*, the most reasonable evaluative approach is to focus on the primary program objective, the improvement of preschool cognitive skills, and to ignore the other potential effects of the program. But this narrower focus undermines the very notion of a comprehensive assessment.

⁴⁴ Note, for instance, that the amount of gain in the ETS evaluations, as discussed in Section III, has not been identified in other than a statistical sense. There is no simple and acceptable measure of the social significance of these cognitive gains.

⁴⁵ It is probably for this reason, rather than the problem of measurement, that benefit/cost analyses of social programs have been so unsatisfactory. For examples of recent analyses, see Joint Economic Committee, *Benefit-Cost Analyses of Federal Programs*, 92d Congress, 2d Session, January 2, 1973, U.S. Government Printing Office, Washington, D.C., 1973.

SUMMARY

This section has attempted to address the two major issues in any new assessment of CTW. The first issue involves the purposes and potential audience for such an assessment. The discussion focused on three major types of evaluations that differed in audience, purpose, cost, and scope. The major choices among audiences appear to be national supporters of children's television programs or the program managers of CTW. The scope of such assessments can cover all of CTW's activities of merely a single activity, and may or may not include deliberate comparisons with non-CTW activities. The discussion did not choose among all these alternatives, but showed how the selection among the alternatives would heavily determine the ultimate evaluation design.

The second issue involves a common assumption made by all of these different types of evaluations. This is the assumption that the full impact of a single CTW activity, such as *Sesame Street*, can be aggregated from a series of individual impact measures. The discussion indicated that such an aggregation is conceptually unsound, and that combined with the methodological problems identified in previous sections, any attempt at aggregating all the positive and negative effects of the activity would be an extremely hazardous venture.

V. CONCLUSIONS AND NEXT STEPS

The Children's Television Workshop (CTW) is engaged in a wide range of activities, dominated by the production of *Sesame Street* and *The Electric Company*. Because of the relative success of these television programs as forms of entertainment and as educational ventures, this study has examined the potential approaches to an assessment of CTW's overall impact on society. The study has explored potential impact measures, reviewed existing sources of information and previous evaluations, and examined the potential framework for an overall assessment, including the use of benefit/cost analysis.

The major conclusion of this study is that:

- *The diverse effects of CTW's activities, indeed even of a single activity like Sesame Street, create insurmountable problems in aggregating the effects in any summary fashion.*

Such a summary statement would not fully account for the true diversity of CTW's impact, and would be a risky procedure that would only draw further controversy. In this regard, benefit/cost analysis is best applied to situations where there exists only one major category of benefits. Any attempt to aggregate several types of benefits, especially if some are positive and others are negative, would be a highly risky procedure.

A second conclusion is independent of the issue of aggregation. It deals primarily with the ability to measure the effects of CTW on a national population:

- *The high audience penetration rates of Sesame Street and The Electric Company will make it difficult to conduct further research on the effects of viewing on a sample of children that is nationally representative.*

Moreover, the existing ETS data for *Sesame Street* only apply to seven specific sites, while the ETS data for *The Electric Company* only apply to two sites. The problem of establishing a national sample does not mean that the effects of viewing as measured by ETS or that might be measured in the future are invalid. The problem does mean that any measured effects cannot easily be generalized to the entire population. Such generalization would be necessary for any statement of the full extent, even on a single measure, of CTW's impact; such generalization is not necessary to establish the internal validity of a small-scale experiment.

These two conclusions lead to a third, which in turn suggests the directions for future steps:

- *The best strategy for assessing CTW's activities is the encouragement of a series of individual studies that will test significant effects singly, and that will demonstrate the true diversity of such effects.*

These studies need not be forced into the same analytic framework, be conducted by the same investigators, or even be sponsored by the same agencies. In fact, the studies would likely call upon a diverse array of research talents. The studies would continue the practice already begun by CTW of encouraging many different inquiries and evaluations, each focusing on a separate but equally important issue. While no precise aggregation of the studies could be made, they would nevertheless in the aggregate constitute a large and incontrovertible body of evidence regarding CTW.

The types of studies to be undertaken should follow from a series of *pre-evaluation* steps: identification of the goals of the various CTW activities, identification of the audience and purpose for any single study, and identification of the broader context within which such a study might be important. To the extent possible, the studies should take advantage of existing sources of financial support in the U.S. Office of Education, the U.S. Office of Child Development, or the foundations, and should not be supported by CTW. In some cases, CTW staff may, however, share as co-investigators on a project.

Based on the issues discussed in this report, the following six studies appear to be the most important and relevant, either to national concerns about educational television or to internal CTW concerns about the effectiveness of its activities, or both:

1. A national survey, conducted by a research rather than a polling organization, that would determine the full size and characteristics of the audiences (target and nontarget) of CTW's programs, especially in comparison to those of other television programs.

The existing audience surveys of CTW's program do not provide accurate assessments of audience size, the duration of viewing (e.g., do most children view *Sesame Street* for one season only?), the amount of adult encouragement for children to view programs, or the extent of nontarget population viewing. A new survey, perhaps coming after the initial season of the new health show, could produce definitive information about the diverse nature of CTW's audience, and compare that audience with those of other public and commercial television programs. Since CTW is primarily a broadcasting organization, measurement of the extent and nature of the viewing audience, regardless of the potential educational effects, is an essential part of any assessment of CTW, and should be given top priority.

2. A multi-year field study, with specially designed experimental and control groups, of cognitive effects of viewing *Sesame Street* and *The Electric Company* alone and in sequential combination.

This study would provide much needed information on the impact of single versus cumulative educational interventions. If single interventions, for instance,

are said to have a longer-term impact only if there also exist interventions in subsequent periods, then viewing of *Sesame Street* for a year followed by viewing of *The Electric Company* for another year should produce larger gains than viewing of either program in isolation. This question of single versus sequentially compounded interventions is an issue that is important not only for educational television, but also for educational programs in general. Such a study would help to test the proposition made earlier, that the benefits from a single intervention would be most appropriately measured in the school year immediately following the intervention, and that the basic rationale for interventions is reliance on the cumulative results of a series of single interventions.

3. A special investigation of the effects of *Sesame Street* and *The Electric Company* on nontarget population audiences; in particular, the effect of *Sesame Street* on slow-learning children and foreign-speaking adults, and the effect of *The Electric Company* on preschool children.

The design of this investigation would follow the general design of the second ETS evaluation, that of selecting specific sites that contain the desired population but that have had minimal exposure to the two programs. While the results would not necessarily be generalizable to a national population, they would fill a large gap in the current knowledge about the unintended benefits from these two programs.

4. An institutional study examining the revenue-raising potential of activities in two or more nonprofit education organizations, including CTW.

This study would attempt to address the major policy issue of the practicality of revenue-raising alternatives to federal or foundation support for educational technology programs (not necessarily limited to television). The study would require the development of new evaluation methods. It would not only have to address revenue-raising potentials, but it would also have to be sensitive to the problem of how revenue-raising activities can be controlled to ensure that the overall nonprofit goals of the organization remain preeminent. For CTW, the conduct of the study would of itself produce much needed information on the economic impact of its activities.

5. A field study to determine the actual amount of in-school effort required to teach the same skills as taught by *Sesame Street* or *The Electric Company*.

This study would provide an initial attempt to determine the social value of specific cognitive gains. It could be combined with an innovative attitude survey of the prices people are willing to pay for various educational gains, resulting from the use of in-school facilities or television programs. The survey would require new types of questions not normally posed on attitude surveys, and should allow for comparisons with the prices people would pay for gains in other public services.⁶⁶

6. A comparative study of the costs and effects of three preschool education pro-

⁶⁶ Jan Acton of The Rand Corporation is currently doing research on such innovative survey techniques, with applications aimed primarily at the prevention and reduction of heart diseases.

grams, *Sesame Street*, *Misterogers' Neighborhood*, and *Captain Kangaroo*, covering both cognitive and noncognitive aspects.

The investigation would attempt to determine the range of effects of these three programs, and would use the opportunity to develop standard measures for comparing educational television programs. The study sample would have to consist of families that had little previous viewing experience with these three programs (e.g., identifying families whose eldest child had just reached the age of two), and have children randomly assigned to viewing and nonviewing (for each of the three programs) over a period of time. The results of the study should be aimed at the resource allocation decisions of the sponsors of these and other children's television programs.

* * * * *

These six studies, then, represent the potential next steps for further assessing CTW's impact. As previously mentioned, each can be undertaken separately under separate sponsorship. The general goal in undertaking all of the studies is to increase knowledge about the diversity of CTW's effects. Even though the information will not be aggregable into a single statement of CTW's impact, such a new array of evidence, combined with the existing evaluations of CTW, will add substantially to our understanding of the impact of CTW on the world.

Appendix

INDIVIDUALS CONSULTED

Dr. Charles L. Bertram
Director of Research and Evaluation
Appalachia Educational Laboratory
Charleston, West Virginia

Dr. Gerry Ann Bogatz
Educational Testing Service
Princeton, New Jersey

Mr. Sol Chafkin
Office of Program Related Investments
Ford Foundation
New York, New York

Dr. Hugh Cline, President
Russell Sage Foundation
New York, New York

Dr. Thomas Cook
Associate Professor of Psychology
Northwestern University
Evanston, Illinois

Dr. George Comstock
The Rand Corporation
(formerly Senior Staff Coordinator,
Surgeon General's Report on
Television and Social Behavior)

Dr. Lois-ellin Datta
Career Education
National Institute of Education
U.S. Department of Health, Education,
and Welfare
(formerly National Coordinator,
Head Start evaluations)

Mr. Paul Firstenberg
Vice President for Finance
Princeton University
Princeton, New Jersey
(formerly at Ford Foundation)

Mr. Clarence Fogelstrom
National Center for Educational
Technology
U.S. Department of Health, Education,
and Welfare
(Program Manager for CTW grant)

Mr. Henry Geller
The Rand Corporation
(formerly General Counsel,
Federal Communications Commission)

Dr. Allen Ginsberg
Office of the Assistant Secretary
for Planning and Evaluation
U.S. Department of Health, Education,
and Welfare

Mr. John Goedtel
Staff member, U.S. House Appropriations
Committee
(member of special study group
examining children's television)

Mr. Joseph Grief
Specialist on Tax Exempt Organizations
Arthur Andersen and Company
Washington, D.C.

Mr. Nicholas Johnson
Commissioner
Federal Communications Commission

Mr. Arthur Kirschenbaum
Office of Planning, Budget, and
Evaluation
U.S. Office of Education

Professor Gerald Lesser
Harvard Graduate School of Education
Cambridge, Massachusetts

Dr. Jack Lyle
Director, Communications Research
Corporation for Public Broadcasting
Washington, D.C.

Dr. J. Ronald Milavsky
Director of Social Research
National Broadcasting Company
New York, New York

Dr. Michael O'Malley
Early Learning
National Institute of Education
U.S. Department of Health, Education,
and Welfare

Dr. Alan Pearce
Research Staff
Federal Communications Commission

Mr. Lee Polk
Children's Programming
American Broadcasting System
New York, New York

Mr. Talton F. Ray
Office of Program Related Investments
Ford Foundation
New York, New York

Ms. Elizabeth Roberts
Children's Television Unit
Federal Communications Commission

Dr. Sally Ryan
Division of Research and Evaluation
U.S. Office of Child Development

Mr. Gerald Sandler
National Institute of Education
U.S. Department of Health, Education,
and Welfare

Ms. Alice Scates
Office of Planning, Budget, and
Evaluation
U.S. Office of Education

Dr. H. Del Schalock
Teaching Research
Oregon State System of Higher Education
Monmouth, Oregon

Mr. Gerry Slater
General Manager
Public Broadcasting Service
Washington, D.C.

Mr. Stuart Sucherman
Public Broadcasting
Ford Foundation
New York, New York

Dr. Joseph Wholey
Director, Evaluation Programs
Urban Institute
Washington, D.C.

Children's Television Workshop:

Franz Allina
Christopher Cerf
David Connell
Michael Dann

Robert Davidson
Evelyn Davis
James Drake
Robert Hatch
Patricia Hayes
Vivian Horner
Thomas Johnston
Thomas Kennedy

William Kobin
Robert Oksner
Edward Palmer
Barbara Frengel Reeves
(formerly on research staff)
Joyce Weil
Norton Wright